

Germanic Obstruent Lenition;
some mutual implications
of theoretical and historical phonology

Patrick Honeybone

NEWCASTLE UNIVERSITY LIBRARY

201 22850 4

Thesis L 7172

A thesis submitted to the Faculty of Arts,
University of Newcastle upon Tyne,
in partial fulfilment of the requirements for the degree of
PhD in Linguistics

2002

	page
Contents	i
Abstract	iv
Acknowledgements	v
1 Introduction	1
1.1 The structure of the thesis	3
1.2 Historical phonology and theoretical phonology	6
1.2.1 Philology and linguistics	8
1.2.2 Theoretical historical phonology	11
1.2.2.1 Phonological processes and phonological change	14
1.2.2.2 Naturalness, phonology and change	17
1.2.2.3 The quanta of phonological change	20
1.2.3 Terminology and concepts in historical phonology	21
1.3 The focus and approach of this thesis	26
1.3.1 Explanation in historical phonology	31
1.4 A first look at lenition	35
1.4.1 Attempts to define lenition	39
1.4.2 The place of lenition in this thesis	43
2 Lenition in Germanic languages	44
2.1 Germanic obstruent lenitions	46
2.1.1 The Germanic Consonant Shift	49
2.1.1.1 The GCS: take 1	50
2.1.1.2 Evidence for the GCS	55
2.1.1.3 Where and when did the GCS occur?	56
2.1.1.4 Is the GCS an example of lenition?	57
2.1.1.5 An appendix: 'GCS' debuccalisation	59
2.1.2 The High German Consonant Shift	60
2.1.2.1 The HGCS: take 1	61
2.1.2.2 Evidence for the HGCS	65
2.1.2.3 Where and when did the HGCS occur?	66
2.1.2.4 Is the HGCS an example of lenition?	69
2.1.3 The English Initial Fricative Voicing	70
2.1.3.1 The EIFV: take 1	71
2.1.3.2 Evidence for the EIFV	73
2.1.3.3 Where and when did the EIFV occur?	75
2.1.3.4 Is the EIFV an example of lenition?	76
2.1.4 The Inner-German Consonant Weakening	77
2.1.4.1 The IGCW: take 1	78
2.1.4.2 Evidence for the IGCW	81
2.1.4.3 Where and when did the IGCW occur?	82
2.1.4.4 Is the IGCW an example of lenition?	82
2.2 Lenition and Germanic: summary and prospect	83
2.3 Reality in reconstruction in historical phonology	84
2.3.1 The use of symbols and their implications	87
2.3.2 The Laryngeal theory	94
2.3.3 The Glottalic theory	97
2.3.4 Summary and prospect	101

	page
3 Theoretical phonology and lenition theory	102
3.1 Theoretical phonology	103
3.1.1 A minimalist basis for phonology	105
3.1.2 Phonological segmenthood	110
3.1.3 Melody: elements and privativity	112
3.1.3.1 Laryngeal specifications	122
3.1.3.1.1 The starting point	124
3.1.3.1.2 What is possible in laryngeal phonology?	125
3.1.3.1.3 What elements are needed for laryngeal specifications? Take 1	127
3.1.3.1.4 Laryngeal Realism 1: fortis/lenis	131
3.1.3.1.5 Laryngeal Realism 2: universalist	135
3.1.3.1.6 What elements are needed for laryngeal specifications? Take 2	138
3.1.3.1.7 Symbols for segments	140
3.1.3.1.8 Laryngeal elements alone: debuccalisation	143
3.1.3.1.9 (De)aspiration and voicing in Germanic languages	147
3.1.3.1.10 Laryngeal specifications in the segments of Proto-Germanic	149
3.1.3.1.11 Laryngeal specifications in fricatives	151
3.1.3.1.12 Summary and prospect	154
3.1.4 Prosody	155
3.1.4.1 Segments in syllables	156
3.1.4.2 The timing tier	158
3.1.4.3 Rhyme Structure: *μμμ	160
3.1.4.4 Summary and prospect	161
3.1.5 Elements in segments	162
3.1.6 Summary and prospect	168
3.1.6.1 Relationships between processes	170
3.1.6.2 Elements and segments: particles and atoms	171
3.2 Another look at lenition	173
3.2.1 Lenition in other languages	175
3.2.1.1 A brief excursus on environments	177
3.2.1.2 Spanish	179
3.2.1.3 Proto-Iranian	184
3.2.1.4 Kannada	185
3.2.1.5 Celtic	187
3.2.1.6 Greek	190
3.2.1.7 Bantu	191
3.2.1.8 Liverpool English	192
3.2.1.9 Summary and prospect	196
3.2.2 What is lenition?	197
3.2.2.1 Lenition observations and generalisations	198
3.2.2.2 Previous accounts of lenition	200
3.2.2.2.1 Lenition scales as explanations	203
3.2.2.2.2 Is lenition just assimilation?	205
3.2.2.2.3 'Ease of articulation'	206
3.2.2.2.3.1 Kirchner (1998, 2000)	209
3.2.2.2.4 Articulatory Phonology	213
3.2.2.2.5 Dependency Phonology	216
3.2.2.2.6 Government Phonology	219
3.2.2.2.6.1 Harris (1990, 1994)	219

	page
3.2.2.3 If these are lenitions, then what is lenition?	222
3.2.2.3.1 Affrication and spread spirantisation	224
3.2.2.3.2 voice stop ‘spirantisation’	228
3.2.2.3.3 Debuccalisation	230
3.2.2.3.4 ‘Voicing’ and delaryngealisation	230
3.2.2.3.5 Summary and prospect	232
3.2.3 Lenition and phonological environments	236
3.2.3.1 Environmental observations and generalisations	237
3.2.3.2 Prosodic lenition inhibition	243
3.2.3.3 Melodic lenition inhibition: melodic molecules	244
3.2.3.4 Summary and prospect	250
4 Historical and theoretical phonology: lenition in Germanic languages	252
4.1 Germanic lenitions and laryngeal specifications	254
4.1.1 The Germanic Consonant Shift revisited	255
4.1.1.1 GCS ₁	257
4.1.1.2 GCS ₂	260
4.1.1.3 GCS ₃	267
4.1.1.4 Summary and appendix: putting the GCS together (with ‘Verner’s Law’)	268
4.1.2 The High German Consonant Shift revisited	272
4.1.2.1 HGCS ₁	273
4.1.2.2 HGCS ₂	280
4.1.3 The English Initial Fricative Voicing revisited	282
4.1.4 The Inner-German Consonant Weakening revisited	285
4.2 The laryngeal history of Present-Day German and Dutch	290
4.2.1 The problem of standards: the case of German	290
4.2.2 Laryngeal switch: the case of Dutch	292
5 Conclusions and mutual implications	298
5.1 Summary of analyses and claims	299
5.2 Philological contributions	302
5.3 Phonological contributions	303
5.3.1 Obstruent lenition	303
5.4 Empirical predictions	305
5.5 Future research and final conclusion	308
References	3010

Abstract

This thesis seeks to reconcile aspects of philology with aspects of linguistics and to illustrate the advances in understanding that can be achieved in this way. The principal empirical focus is provided by four sets of ‘classic’ data from Germanic historical phonology and the theoretical background is provided by recent developments in non-linear models of phonology. The mutual implications of these are addressed to provide both a new understanding of the data and a coherent theoretical understanding of the types of phonological process that they exemplify. The data sets are chosen partly because of their iconic status in historical phonology and partly also because they all, at times, have been described as examples of a general process-type: ‘lenition’.

The background is provided by chapters one and two. Chapter one introduces the way in which the relationship between historical and synchronic phonology is understood in the thesis. Chapter two introduces the classic sets of data, namely: (i) the ‘Germanic Consonant Shift’, (ii) the ‘High German Consonant Shift’, (iii) the ‘English Initial Fricative Voicing’, (iv) the ‘Inner-German Consonant Weakening’.

Chapter three problematises these sets of data in two main ways. The first of these is through an investigation of how they fit with recent advances in the understanding of phonological structure and phonological processes which have been developed in phonological traditions that see subsegmental units as privative (eg. Dependency Phonology, Government Phonology, Articulatory Phonology). Particular emphasis is placed on the units needed to account for laryngeal specifications (traditionally described as ‘aspirated’, ‘glottalised’, ‘voiced’ and ‘voiceless’) and a position is defended whereby three privative units are required: |spread|, |voice| and |constricted|. The first two of these are shown to be used differently in languages to account for the two traditional categories of ‘voiced’ and ‘voiceless’.

The second problematising factor is a discussion of how the data sets fit with the concept of obstruent ‘lenition’, a notion which has a clear, if problematic, place in both historical and synchronic phonology. Lenition trajectories of the type ‘stop > affricate > fricative > glottal’ and ‘voiceless > voiced’ are discussed in the light of the data introduced in chapter two and the theoretical discussion of subsegmental structure developed in the first half of chapter three. Further data from Dutch, Spanish, Dravidian, Indo-Iranian, Greek, Bantu, Celtic and Liverpool English is discussed and a synthesis is proposed which takes into account both attested historical data and a theoretical phonological understanding. A notion of ‘lenition inhibition’ is developed to explain some of the exceptions (which are frequent in lenitions, but are rarely discussed) whereby the sharing of autosegmental phonological units gives a segment ‘strength’ in certain environments.

Chapter four revisits the data introduced in chapter two and provides a reinterpretation in accordance with the understanding of ‘lenition’ processes developed in chapter three. The data in (i) and (ii) are essentially shown to be dependent on the presence of a |spread| laryngeal specification and to be qualitatively different from the data in (iii) and (iv), which are shown to be unifiable as the loss of |spread|. Additionally, the patterns of exceptions to these processes are examined in light of the understanding of lenition inhibition developed in chapter three.

In chapter five, general conclusions are drawn. It is argued that not all lenition processes are unifiable as a single process type. It is further shown that previous discussions of historical ‘lenition’ processes have missed generalisations, both in terms of the necessary conditions for their diachronic innovation and in terms of the environments which can be seen to partially inhibit their introduction. It is also shown how a detailed understanding of the diachronic data which is discussed here can make a contribution to theoretical models of phonology.

Acknowledgements

I owe the greatest academic debt of thanks to Phil Carr, who supervised me initially at Newcastle and has still supported me since he moved to Montpellier and I moved to Edge Hill. I owe my understanding of phonology and phonological thinking to him. I would also like to thank others at Newcastle University who taught me and talked with me over the years I spent there. Particularly Noel Burton-Roberts in the English department and Jon West in the German department have helped to open my eyes to the wonders of general and Germanic linguistics. Thanks are also due to Joan Beal, who took me over as a supervisee, and to Geoff Poole, who helped me through the last administrative hurdles. I'd also like to thank those who taught me at Würzburg university. I think it was there that I first understood how amazing historical phonology can be.

I would also very much like to thank those who have supported me in many ways while I have been writing and thinking about this. My parents have always done all they could to encourage and support me throughout my life. And much love and many thanks are due to Sharon Unsworth and to Liz and Tony, Geoff Poole (again), Maria Maza and Nils Langer. The last two especially have also been worthy companions in linguistics.

Thanks are also due to certain definite parts of the Edge Hill community for support, friendship and for working to ensure that research can continue there, sometimes against the odds. Joy Bowes, especially has helped make it possible for me to write this. I'm also grateful that I've been able to discuss phonetics, phonology and lenition there with Deborah Chirrey, Clive Grey and Kevin Watson. Latterly, also, Liverpool University Library deserves some kind of acknowledgement for giving me a home among the ancient Germanic Grammars and the linguistics section, unbeknown to them, even though it was.

Many other people have also discussed various phonological issues with me, or have sent me their work, and I am grateful to have had the opportunity for such interaction with Ricardo Bermúdez-Otero, Wiebke Brockhaus, Mohamed Elmedlaoui, Michael Jessen, Lisa Lavoie, Tobias Scheer, Michael Wagner and many others, although none of those named here are in any way responsible for the use that I've put their work to, and I'm sure that they'd not agree with all or even any of my conclusions. Finally, it will be clear that I owe a substantial debt to many of those whose work I have cited, although they, too, might not agree with what I've done with it.

I was supported financially while writing and working on this by a full PhD studentship from the British Academy (which turned into the Arts and Humanities Research Board half way through) for which I am also very grateful.

NB: This thesis was completed and submitted in May 2001; it and I were examined by April McMahon and Carol Fehringer in July 2001, passing with minor typographical corrections, so that is the true date of its completion. Due to circumstances beyond my control, however, I was not allowed to graduate until 2002.

*Alles Gescheide
ist schon gedacht worden,
man muss nur versuchen es noch einmal zu denken*

*Everything worth thinking
has already been thought already,
you just need to try to think it again*

Johann Goethe

*Kein Geringerer
als Leonardo da Vinci
lehrt uns
»Wer immer nur Autoritäten zitiert
macht zwar von seinem Gedächtnis Gebrauch
doch nicht
von seinem Verstand«
Prägt euch das endlich ein:
Mit Leonardo
los von den Autoritäten!*

*No less a personage
than Leonardo da Vinci
instructs us
"Those who simply quote from authorities
may well be using their memory -
but not
their intelligence"
Will you get this into your heads:
follow Leonardo
forget your authorities!*

Erich Fried (1978, 100)

*Uebergang stimmloser Verschlusslaute in stimmlose Spranten ist sehr häufig in den
verschiedensten Sprachen, doch ist seine Entstehung nicht überall mit Sicherheit festzustellen...*

Die p, t, k werden aber in verschiedenen Gegenden ganz verschieden ausgesprochen...

*The change from voiceless stops into voiceless fricatives is very common in a wide
range of languages, but the genesis of the change is not always entirely clear...*

However, p, t, k are pronounced very differently in different areas...

Sievers (1901, 285 & 144)

1 Introduction

If it searched hard enough, linguistics could probably find its foundations in historical phonology. Or, if that is a little overblown, then it is true to say that the systematic method which was developed for the study of sound change has had profound effects on the study of language. It might have to share the glory with historical morphology, but the early type of diachronic and comparative phonology that is associated with the now canonised names of Rask, Grimm, Verner, and the neogrammarians can claim a place as a founding aspect of the basic types of modern linguistics. The gradual realisation of regularity which occurred through the nineteenth century led to the neogrammarians' 'exceptionless hypothesis' in order to account precisely for phonological processes that had been innovated in languages. This was probably the first predictive principle in linguistics, and it has since been echoed in the notions of phonological rules, syntactic principles and strict dominance in constraint hierarchies.

Modern theoretical phonology seeks to deal with exceptionlessness, too. Predictions are made about what is a possible linguistic system and what is a possible phonological process. These predictions result from the models that phonologists have built up to account for precisely the kind of phonological patterns that were first noted by early historical and comparative phonologists. In this thesis I seek to combine aspects of modern phonological theory with some of the crucial foundational observations that have been made in historical study. I argue that, even after their long common history and a notable tradition of interaction which has continued since they became recognisably independent disciplines, historical phonology and theoretical phonology can still find interesting implications for each other, even by examining some of their most basic assumptions.

In the development of linguistics from the quasi-mystical start that I have just given it,¹ it was natural that different strands of linguistic study should develop, and it is entirely obvious now that a resolutely synchronic theoretical phonology stands opposite a branch of historical study which has not stood still since the neogrammarians, but has concentrated on developing methods devoted to working out the history of individual languages. The two have always interacted, but it is now possible to identify distinct disciplines which should, I believe, interact more than they do: (i) a strand of 'theoretical phonology', which focuses on questions such as what is possible in phonology and how the analyses that we make of pieces of data from

¹ For a more realistic, but I think fundamentally compatible account, see for example Arens (1969), Robins (1990) and Davies (1998); and see Kiparsky (1988) and Bromberger & Halle (1989) for some comments on the influence of historical phonology on theoretical phonology and of theoretical phonology on syntax.

individual languages can help us to understand the general patterns that are found again and again in phonological systems, and (ii) a strand which focuses on the history of particular linguistic systems, bringing all available evidence to bear (manuscripts, orthoepic descriptions, dialectology and the like) in an attempt to understand the fundamental patterns of change that have occurred in a particular language.

The title of this thesis presupposes that there are these identifiable and discrete entities. It further presupposes that we can name them ‘theoretical phonology’ and ‘historical phonology’ (but as I shall show below, these names are not really straightforward) and that we might hope for novel effects when we bring them together. If it is successful, this thesis will make a contribution to both theoretical and historical phonology. It seeks to do this by exploring and illustrating the relationship that exist between the two. At times we will see that differing terminological traditions, even the different use of symbols for transcription in the two disciplines can lead to misinterpretation. I will, in fact, claim that mainstream synchronic phonological theory has at times sorely mistranscribed certain segments and hence has missed important generalisations, both in terms of how contemporary languages behave and in terms of what kinds of phonological process are possible. The insight for this realisation derives from work in the historical tradition (although even this tradition has not consistently applied it cross-linguistically). This illustrates the importance of being open to mutual implications.

Some of the key data in this thesis can be traced back to almost the earliest discussion in historical work on Germanic phonology. It may perhaps seem a little foolhardy to deal with such data because of its founding glory for the discipline and the wide attention that it has already attracted. It is, however, just data like any other, and I argue here that it, too, has been misanalysed up till now, or rather, the conclusions that have been available to scholars in the area have not been consistently drawn, and have certainly not been drawn with the insight that contemporary models of theoretical phonology can bring to the questions.² I also investigate other, less well-known data in the same light, that is, considering what we know about the general kind of processes that they exemplify.

The data, as can further be gleaned from the title to the thesis, is principally of Germanic origin. The four key sets of data which I investigate and analyse are taken from the history of several Germanic languages. They are not considered on their own, however, and are overtly brought into a broader picture of universalist phonology. The thesis’s title also promises

² In dealing with such data, we can also agree with Vennemann (1983) that “there are no better testing grounds for new theoretical proposals within a science than its oldest and most elaborated problems” (1983, 6).

‘lenition’, and this derives from the fact that (at least most of) the phonological processes which are involved in the four main sets of data have, at one stage or another, been described as examples of lenition, an overarching process-type of a sometimes uncertain status. A further key goal of this thesis, after and alongside the investigation of the specific pieces of data for Germanic is to dissect the notion of lenition and to ascertain whether it has any explanatory potential, to consider to what extent it helps us to understand the processes that can be seen to have occurred in the Germanic data, and indeed to understand analogous processes in any other language.

The ‘lenition processes’ which I investigate are exclusively ones which affect (and mostly produce) obstruents. As we will see, the term ‘lenition’ has been used in some previous work with a wider frame of reference than that, but it would not be feasible in this thesis to consider every type of analysis that has ever been described by the term. With the restriction in the area of study to the closely knit natural class of obstruents we might well think that if the notion ‘lenition’ can be shown to furnish us with insightful analytic tools for the discussion of the diachrony of obstruents, then these may, subsequently, be transferable to improve the investigability of lenition in other types of segments. If, on the other hand, lenition fails as an explanatory device after just the discussion of obstruents, then the chances that it will help us to understand vocalic developments, or those in non-obstruent consonants, are minimal.

1.1 The structure of the thesis

In this thesis I investigate arguments and material from a number of academic traditions; one of the aims of this thesis is to explore the links between them. As we saw above, work has been done for centuries in some of these traditions, and some of the data that I deal with here has long featured in discussion in these traditions. It will therefore not be possible to provide a completely comprehensive ‘literature review’ on all the topics discussed. A vast amount has been written on some of the key theoretical points and on some of the data which is discussed here; a large amount has been written on the notion of ‘lenition’ and an even larger amount has been produced on the nature of historical phonology and on theoretical approaches to the understanding of phonology in general, and I do not review it all.

Nonetheless, a good deal of existing literature on the topics discussed here is dissected, both recent and ‘classical’ works. This typically occurs in the place where it makes sense to discuss it, so some of chapters 1, 2 and 3 is devoted to a review of previous proposals. At other points

in those chapters, however, especially in 3, but also in 2, novel theoretical and analytical positions are proposed and defended. Chapter 2 chiefly consists of a presentation of the key Germanic data that is discussed in this thesis, but it also deals with some important matters of interpretation. I return to the data in chapter 4, to present a set of original analyses; these analyses rely on positions and proposals about the nature and specifics of phonology which are explained in chapter 3. Chapter 1 is largely introductory and chapter 5 conclusory. Certain further specifics regarding the contents of the individual chapters are given below.

The rest of chapter 1 explains some of the key assumptions which I make in the thesis and explains what I mean by certain key terms that I use in later chapters (for example the terms ‘historical phonology’ and ‘theoretical phonology’ are addressed more closely in section 1.2). This will partly involve a gradual focusing-in on the areas which will be the subject of discussion in the remainder of thesis, and this will naturally mean that certain important aspects of phonological and historical interest must be set aside as it is recognised that, while they are important areas of study in their own right, they will not prove important here. In a rather brief fashion, a series of distinctions will be established which will serve to split up the wider area of study into coherent parts, some of which can be set aside in a justified manner, to allow a focus on a manageable topic. The chapter is also intended to provide a basic understanding of certain ideas which will allow us to proceed to the discussion of the data in chapter 2, but which we will return to later in the thesis, where they will be shown to be over-simplistic. This is the case for the initial discussion of the notion of lenition, which occurs here in section 1.4. The chapter also provides a discussion of certain important notions which we will return to later, including the place of ‘naturalness’ in analysis (section 1.2.2.2), the extent to which we can claim to provide explanations for historical data (section 1.3.1), and the relationship between historical and synchronic data in phonology (section 1.2.2.1).

Chapter 2 discusses four sets of data which comprise changes or sets of changes in the history of Germanic languages. These are dealt with under the names (i) the Germanic Consonant Shift, (ii) the High German Consonant Shift, (iii) the English Initial Fricative Voicing and (iv) the Inner-German Consonant Weakening. As we will see, these all feature phonological processes which have been described as lenitions. The discussion there involves an investigation of our knowledge of the data; this deals with such questions of how secure this knowledge is, how best it can be interpreted, and a detailed discussion of the form in which the

data is presented and understood. The chapter concludes with a section which confronts traditional analyses of the sets of data with universalist phonological analysis.

Chapter 3 is the largest chapter, because it consists of two half-chapters which deal with slightly different issues, but which have a clear thematic unity. They problematise and rationalise the data presented in chapter 2 in terms of, (i) a detailed investigation of aspects of contemporary phonological theory and (ii) an investigation of approaches to the understanding of lenition. The discussion of (i) occurs in section 3.1, and focuses particularly on aspects of melodic, or segmental representation. The most substantial discussion is devoted to laryngeal specifications, as these will prove most important in the discussion. Certain features of prosody are also discussed. The discussion of (ii) occurs in section 3.2. This features first a short discussion of a number of attested phonological processes which are analogous in certain ways to the Germanic lenitions discussed in chapter 2, widening the empirical base to enable broader generalisations about the nature of lenition. This section then investigates a number of previous approaches to lenition, before presenting a novel understanding of the processes involved and a novel approach to the interaction between lenition processes and phonological environment. The chapter recognises certain generalisations about possible lenition processes and environments which have not been noted previously.

Chapter 4 returns to the sets of Germanic data which were first introduced in chapter 2 and presents a reanalysis of them which is consistent with the understanding of phonology in general and lenition processes in particular which is developed in chapter 3. Chapter 4 also features a discussion of certain other aspects of the phonology of Germanic languages which are clearly affected by the phonological analyses that are presented in chapter 3.

Chapter 5 summarises the conclusions which are drawn and the claims which are made in the thesis and examines certain final implications of them. Their empirical predictions are drawn out and certain cases of potential counterevidence are dealt with. Finally, some pointers are given to indicate how future research can help to examine these implications and to test the predictions that are made in the thesis.

The thesis thus includes certain big claims and some rather smaller claims. The smaller claims are the analyses of the Germanic data in chapter 4; the larger claims reside in the model of phonology and the understanding of lenition which provides the background to these analyses. The claims are not all dependent on each other, in that some can fall while others stand, but they are all connected, as we will see. They all conspire, I propose, to improve our understanding of the data that I present in chapter 2. In this thesis, I naturally make use of

certain ideas and positions which have been developed by others to deal with certain types of data. Part of the novelty here lies in the application of these ideas to types of data which they were not designed to account for; for example, I apply a certain set of ideas which have been developed to account for synchronic laryngeal phonology to the study of diachronic phonological change. Certain other more fundamental claims are made here as well, however, and these contribute to the novelty of the work in the thesis.

The novelty of some of the claims and the application of certain recent theoretical ideas to new sets of data in the thesis mean that it deals with points which are relevant to contemporary debates in the fields to which it seeks to contribute. However, it might be contended that this thesis also contains discussion of concerns which seem anachronistic to the 21st century phonologist. There is no discussion here which will contribute to issues within Optimality Theory, for example, or the Strict CV model of ‘syllabic’ structure. Rather, I take the opportunity to address some fundamental issues which have a long past. This is partly because I believe they will also have a long future and will not become the ‘rule ordering’ debates of the hereafter. It is also because, even if they are not burning topics of debate at the moment, I believe they should be remembered and returned to.

The quotations at the very start of this thesis are meant to imply that great work has been done on the topics that I address here and that some of the positions which I propose are in fact a revisitation of older insights through new ways of understanding. The quotations are also meant to imply that great and copious previous work can also be wrong. We need to have a certain respect for what has come before (and hopefully an awareness of it, so that we avoid the danger of reinventing the wheel), but it must be a critical and questioning respect.

1.2 Historical phonology and theoretical phonology

Because the terms ‘historical phonology’ and ‘theoretical phonology’ are given a clear prominence in the title of this thesis, it will be well to devote some space to explaining what I mean by them and how this fits with other usages. It is, in fact, not entirely straightforward to draw a simple distinction between the two. The basic distinction that I mean by the contrast is that ‘theoretical phonology’ seeks primarily to devise a detailed understanding of the phonological aspect of the language faculty and ‘historical phonology’ seeks primarily to understand the patterns and details that can be discovered in diachronic change in the phonology of languages. But it is not really so simple to make the distinction. One relatively

clear reason for this difficulty in differentiation is that there is a substantial body of work which overtly works to combine the two and could easily be described as ‘theoretical historical phonology’. Much work that could be described in this way has naturally informed the work undertaken in this thesis.

Another reason why perhaps the distinction is not simply made is that there is a fair amount of work in what is principally ‘theoretical phonology’ (because it is primarily intended as a contribution to the understanding of synchronic models of phonology) which makes use of both contemporary and historical data, and sometimes even of both synchronic and diachronic data. For example, Bromberger & Halle (1989), in a paper which they see as setting out some fundamentals for a theoretical model of phonology, happily analyse the Germanic Consonant Shift and Verner’s Law (which, as we shall see in chapter 2, are separated from the present by millennia) along with contemporary phonological phenomena. Several introductory textbooks on phonological theory feature sections on historical phonology (for example, Hyman 1975 and Lass 1984) and even the foundational text for much of theoretical phonology, Chomsky & Halle (1968), has a chapter on historical issues.

Thus a fair amount of work which we might want to describe as ‘theoretical phonology’ deals in historical data. However, the most common use for such historical data in theoretical phonology is not truly diachronic, but as examples of synchronic processes which just happen to have been phonologically active in the past. This is an entirely reasonable use of such data, indeed it illustrates the crucial underlying principle of historical linguistics, the uniformitarian principle (which Kiparsky 1988 attributes to Scherer 1868 Chapter and the Neogrammarians), and which simply states that the human languages have always obeyed the same principles,³ which means that whatever we observe happening today is relevant to help explain the phonology of past linguistic states, and that what we can learn from past phonological states can help us to understand the phonology of Present-Day languages. While it is the key assumption without which there could be no historical linguistics,⁴ and is generally taken as read, I will have occasion to refer to the uniformitarian principle in argumentation below, which is why I make mention of it here.

³ The ‘always’ here naturally only goes as far back as the evolution of the human language faculty to its contemporary state, but we have no hope of reconstructing protolanguages as far back as that or further, so the statement is true enough as it is formulated. See for example Salmons (1992) for discussion of the limits of reconstruction and McMahon (2000b) for some points relevant to the evolution of phonology.

⁴ See Lass (1997) for some detailed discussion of the idea and the issues which are connected with it. Lass notes that the concept is also referred to as ‘actualism’, but I retain the traditional term.

The principles of phonology are what theoretical phonology seeks to make explicit, and many historical phonologists seek to contribute to this goal, arguing that diachronic and other historical data can help us to understand what is possible in phonology and how we can expect phonology to operate (for example, Kiparsky 1968, 1981, King 1969, Lass 1976, Iverson & Salmons 1995, 1999a, Bermúdez-Otero 1999, McMahon 2000a). This could be a description of the ‘theoretical historical phonology’ which was mentioned above. I return to discuss certain aspects of this work and how the discussion in this thesis fits in with it below, but first I turn to another possible understanding of the phrase ‘historical phonology’. This focuses exclusively on diachronic concerns and brings us back to some of the foundational work mentioned at the very start of this introduction.

1.2.1 Philology and linguistics

A related contrast to that between ‘theoretical phonology’ and ‘historical phonology’ is the distinction between ‘linguistics’ and ‘philology’. In the abstract to this thesis and in discussion below, I speak of a distinction between these two academic disciplines and of my hope to combine aspects of the two. A distinction along these lines can be made, but on close inspection it, too, is far from straightforward (as is normally the case with such distinctions). One problem with making the distinction is the slipperiness of the meaning of the term ‘philology’. Campbell (1998) gives three understandings of the term:

Sometimes philology is taken to be merely the study of some classical or older language.... Sometimes philology is understood to mean historical linguistics as practised in the nineteenth century.... In another sense of the word, philology is understood as the scholarly activity which attempts to get systematic information about a language from written records.

Campbell (1998, 328-9).

I mean the term to refer to something which is largely an amalgam of all the meanings that Campbell gives. There is a clear tradition of scholarly research into the history of languages which has historically focused on ‘older’ stages of languages and which commenced early in the nineteenth century. This developed through a series of extraordinary achievements to obtain real sophistication during that century.⁵ It is also notable that much of this work was

⁵ As is well known, the unbroken line of resolutely historical texts which form the nineteenth century canon of linguistic achievement include Bopp (1816), Rask (1818), Grimm (1822-37), Schleicher (1861), Verner (1876), Brugmann & Delbrück (1886-1900). These illustrate a gradual break from the chiefly philosophical bent of the

conducted either in German or by Germanicists, and this is reflected in the references (and elsewhere) in this thesis.

This tradition, which we might reasonably call a ‘discipline’, or perhaps even a ‘research programme’, has accumulated a phenomenal wealth of data regarding the diachronic phonology (and morphology, but not so much syntax) of a wide range of languages (although especially those from the Indo-European family) which is deposited in works which are often referred to as ‘the Handbooks’ (see, for example, Lehmann 1993 for a discussion on the term ‘Handbooks’).⁶

In common with other similar work, I refer to ‘the Handbooks’ in what follows, particularly in chapter 2. These Handbooks might almost be taken as another way of defining what is meant here by philology in that they can be seen as a crowning glory of the philological approach and are a remarkable storehouse of knowledge about ancient synchronic states of languages and of descriptive accounts of their diachronic development. I take these to include the ‘standard Handbooks for Indo-European’ such as Brugmann & Delbrück (1886-1900) and Meillet (1937)⁷ and also the historical ‘Grammars’ included in such series as the *Sammlung kurzer Grammatiken germanischer Dialekte* (the ‘collection of short grammars of Germanic dialects’) such as Braune’s Grammars of Gothic and Old High German, Paul’s for Middle High German, and Sievers’s for Old English, which were typically first published in the late nineteenth century and have since gone through a number of editions, which have seen them being updated with reference to newly appeared literature but often largely maintaining the general approach of the earliest editions. These Handbooks are my chief source for the data which is discussed in chapter 2, especially, of course, the Germanic Grammars.

I refer to this tradition as ‘philology’. It focuses on diachrony and often takes as a key aim the understanding of the precise nature of the history of individual languages. Given this, it is reasonable to refer to the work which has been done on the phonology of languages in this tradition as ‘historical phonology’, that is, simply, the study of the diachronic events in the phonology of languages. This usage of ‘historical phonology’ is rather different from the usage

types of language study which directly preceded them, and the development of a systematic approach to the study of linguistic form, isolated from the social or literary context in which it occurs.

⁶ It’s not completely clear that we would be justified in using the term ‘research programme’ given the range of theoretical backgrounds which were assumed by different writers and at different periods in the century. Nonetheless, with hindsight it is not difficult to see the work of the nineteenth century as leading up to the production of the Handbooks although, really, this would impute too much teleology to the development.

⁷ See Lehmann (1993, especially chapter 3) for a lucid discussion of these Handbooks, their production and achievements, along with the discussion in Robins (1990), Fox (1995) and much other work on the history of linguistics.

that might be connected with the work of Kiparsky and others which was briefly mentioned above, and this is why the term is somewhat ambiguous: does it refer to the cataloguing of the diachronic phonological events in the historical of languages or to attempts to understand the principles of phonology using the evidence of such events? In the title of this thesis, both meanings are meant. There are two key sources of evidence for philological research into earlier states of languages and their diachronic development: (i) the careful consideration of written records and (ii) the comparative method of reconstruction. Some of the discussion in this thesis connects with concerns which are relevant to both of these research tools.

If we define 'philology' as I have just done, it is possible to contrast it with 'linguistics' (part of which is 'theoretical phonology'). It will be clear, though, that 'philology' is by no means entirely unconnected from 'linguistics'. This is at least in part because modern structuralist linguistics has partly grown out of this nineteenth century philological tradition, as was mentioned at the very start of this chapter. At the same time, though, 'philology' is a contemporary of 'linguistics', because such purely historical work continues today, of course, and could entirely reasonably claim to be a branch of 'linguistics' itself. So the only way to contrast 'philology' and 'linguistics' is to restrict the latter to the discipline which might also be referred to as 'theoretical linguistics' and which has largely turned its back on diachronic data to focus on the establishment of an explanatory account of the human language faculty.

The difference between the two disciplines is clear from the fact that the asterisk '*' is used so differently in them: to mark a reconstructed form in philology, and to mark an ungrammatical form in linguistics. At least in part because both disciplines have influenced the work in this thesis, and so there is potential for confusion, I do not make use of the asterisk at all here. It will be clear from the discussion where the phonological segments which I refer to are reconstructed, and where something is impossible in phonology, I make this explicit.

The above terminological tangle is not helped by the existence of the discipline of 'historical linguistics'. This can probably be recognised as an area of study, of which 'theoretical historical phonology' may be a part, which seeks to combine aspects of philology with aspects of linguistics. This thesis tries to situate itself here. This means that it is important to focus on philological detail, but also on linguistic theory. I seek to combine theoretical insights into what is possible in language from (theoretical) linguistics with the wealth of diachronic data which has been gathered in (historical) philology. Perhaps because of their subject matter, both philology and historical linguistics have an acute awareness of their own history as disciplines.

At places this is reflected in the approach taken in this thesis; this also reflects the comments at the end of section 1.1.

In combining the two potentially separate areas, both can benefit from a realisation of their mutual implications. Theoretical models of phonology can be tested or refined by the confrontation with historical data, and philological data can be revisited according to what we know about language in general, and may be re-interpreted in this light. This thesis aims to do both of these. In the next sections, I discuss certain key points which are relevant to this kind of study. I turn first to some concerns of theoretical historical phonology and then to points which are relevant to general historical phonology

1.2.2 Theoretical historical phonology

At the start of this section (1.2), I described a ‘theoretical historical phonology’. This is a tradition which has close links with theoretical models of phonology, and this means that as phonological theory has developed, so has theoretical historical phonology. There has been historical work in such theoretical frameworks as Standard Generative Phonology (for the framework, see for example, Chomsky & Halle 1968), Lexical Phonology (eg, Kiparsky 1982b), Dependency Phonology (eg, Anderson & Ewen 1987), Government Phonology (eg, Kaye, Lowenstamm & Vergnaud 1985) and Optimality Theory (eg, Prince & Smolensky 1993) and some of this work finds an echo in this thesis. It is natural that the greatest influence comes from only some of these models and that the concerns of the others will not be addressed. The model of phonology which I adopt in section 3.1 is most heavily influenced by theoretical frameworks such as Government Phonology and Dependency Phonology, which adopt a distinctive approach to segmental structure using only privative units (the rationale behind this is explained in section 3.1.3). Other aspects of the model that I adopt are also explained in section 3.1, so I will not discuss them in detail here, but a few points will help to explain to general approach to be taken.

As is well recognised (see, for example, Anderson 1985, Ewen & van der Hulst 2001), it is possible to divide phonology into theories of representations, derivations and levels. The key phonological focus in this thesis is on phonological representation. As implied above, I discuss a theory of representations in section 3.1. In fact, I argue that certain aspects of historical data have implications for the model of segmental structure that we should adopt. I also make some proposals connected with a theory of levels and of derivations, but these are less important

here. To an extent, the nature of the theoretical machinery that is adopted to map representation onto representation is not so important for the work in this thesis, and I do not discuss this in any detail. All of the phonological processes which I discuss have an exclusively phonological or phonetic motivation and conditioning and I do not deal with processes which models such as Lexical Phonology would view as being situated at a low-numbered lexical stratum. The processes which we encounter in this thesis would be classed in such models as either 'postlexical' or in a high-numbered stratum in the 'lexical' phonology. As will become apparent in section 3.1, the model of phonology which I adopt has no place for such processes as 'vowel shift' in English (as Chomsky & Halle 1968, Halle & Mohanen 1985 and others have analysed it); I see such phenomena as diachronic phonological events, which are now lexicalised into underlying representations. Much of what I propose and discuss here can be considered separately from considerations of derivation and levels, however, and to a fair extent it could be quite straightforwardly recast in the theoretical clothes of other frameworks.⁸

There are certain fundamental phonological concerns which all models share, however they may be formalised, and these are naturally transferred to theoretical historical phonology. One leitmotif of much of phonology is the search for phonological universals. This has long had a place at the heart of phonology, and there have been many attempts to define what is a possible phonological process and what is imaginable, but never occurs.⁹ This also delimits what is a possible phonological change, to the extent that phonological change is process innovation. Hyman (1975) illustrates this in a now classic introductory textbook to phonology (this also illustrates the point made at the start of section 1.2 that many phonologists who chiefly see themselves as synchronic linguists also make reference to diachronic data). Hyman writes:

While it is a well-known fact that sounds change through time, some sound changes are more frequently attested than others, while still other potential sound changes are not attested at all. For example, the sound change turning [b, d, g] into [p, t, k] has been observed in several languages families of the world (for example in the history of

⁸ I discuss certain aspects of the translatability of phonological analyses from one framework to another in section 3.1. In general, it seems clear that much work in many frameworks is not difficult to reinterpret into the terminology and constructs of another theoretical model. It is important to view data from the perspective of a theoretical framework, as this can help to frame hypotheses as to how we might expect the data to be interpretable, but it is also important to maintain a wider picture of what all types of phonology have in common. It is naturally not the case that all ideas in all frameworks can be right, as some are indeed incompatible. For example, while the analyses which are presented here in 'process' terminology could be recast in terms of an Optimality Theoretic constraint ranking, it is not clear that they fit with the common OT assumption of the 'Richness of the Base' (for discussion of this, see section 3.1.2).

⁹ The existence of phonological universals has taken on an increasing importance in recent years (encouraged by the rise of Optimality Theory, see Prince & Smolensky 1993, McCarthy & Prince 1995).

Chinese). This change constitutes part of the consonant shift known as Grimm's Law, which separates the Germanic branch from the rest of the Indo-European languages. On the other hand, a sound change turning all instances of [p, t, k] into [b, d, g] has never been reported. If such a sound change were to take place, the resulting system would include a series of voiced stops but no series of voiceless stops. In other words, the Jakobsonian implicational universal whereby /b, d, g/ implies /p, t, k/ would be violated. As pointed out by Greenberg [(1966, 510)], any sound change which produces an impossible sound system (such as the one which would result from a change voicing all voiceless stops) is an impossible change.

(Hyman 1975, 17-18)

This passage is worth quoting in full because it eloquently expresses the perspective in question, but also because it illustrates the omnipresence of some of the key data discussed in this thesis ('Grimm's Law' - first discussed here in section 2.1.1) and also because it makes a claim about possible process types, which would probably find general agreement among phonologists. This is interesting here because, as we will see in section 2.1.4, the claim seems to be falsified by one of the sets of Germanic data which is discussed in this thesis. One of the key points made in later chapters, however, is that while such seemingly falsificatory data exists, in fact it only *appears* to falsify the universal which Hyman (1975) mentions because the conventional analysis of it depends on a faulty theoretical model. This will rescue a form of Hyman's generalisation and back up the basic point.

The terminology that I adopt in this thesis is slightly different to that which Hyman uses above. As is quite normal, Hyman speaks of 'sound change' and of 'change' in general. In what follows, I typically refer to 'phonological change' because any change in the pronunciation habits of a group of speakers, in the sounds that they produce, is a reflection of a change in their phonology. This may simply involve a clearly phonetically motivated process which has been captured by the phonology of the speaker as an only slight change in the surface correspondence to an underlying form, but this is nonetheless still a change in the phonology of the variety. Phonological change can thus involve the addition of a phonological process to a speaker's mental phonology, and this might have the effect of altering the nature (and number) of surface segments in a variety; other types of phonological change might involve change in the nature of underlying segments, and this might lead to a reduction or augmentation in the nature of possible contrasts in a variety.

This latter type of change has at times been the main focus of phonological attention, to the extent that change in 'predictable properties', such as might be produced by a phonological process, has not been seen as particularly important. 'Structuralist' phonemic approaches

which focussed chiefly on segmental merger and split (see Hoenigswald 1960, for example), would not necessarily count some of the diachronic phonological events that I focus on here as ‘important’ because they did not alter the number of contrasts. While it is true that changes in contrasting segments can have an important effect on a phonology, I maintain that any change in the phonology of a variety is worthy of study as a phonological change, even if it only altered certain ‘allophones’ of an underlying segment or the pronunciation of all occurrences of an underlying segment, without affecting its contrastiveness. In some of the discussion to follow, the phonological notions of contrast and segmental merger will be an important feature of discussion, especially in section 2.1.4 (in chapter 2) and section 4.1.4 (in chapter 4), but they do not guide the whole focus of the thesis by any means.

I focus here quite overtly on the diachronic innovation of certain phonological processes some of which have not received the attention that they deserve. This is at least in part because some of them did not affect the underlying phonological system of contrasts (that is, the phonological inventory of a variety). They did, however, by the very fact of their innovation, affect the variety’s phonology, and their innovation is potentially as interesting as the innovation of any phonological process. An understanding of the means by which such phonological change can be innovated will form an important part of the picture here, and I turn to this in the next section.

1.2.2.1 Phonological processes and phonological change

It is sometimes argued in historical linguistics that the key locus for linguistic change lies in acquisition. This was already present in Paul (1886) and also features in much contemporary discussion (see, for example, Lightfoot 1991, 1999). Nonetheless, it has also been recognised that change can occur in speaker’s phonologies during their lifetime. In this thesis, I place more emphasis on this latter locus of change than is sometimes the case in discussions of historical linguistics.

It clearly is the case that the process of language acquisition can have a vital role in phonological change. A language has a chance to start anew in acquisition (to overpersonify the situation) and in this way, the acquisition of the phonology of a language by a new generation can lead to ‘catastrophic’ changes, or at least substantial differences, if a previously existing situation is ‘reanalysed’ by a new generation (it might be better to say ‘differently analysed’ because the acquirers involved never analysed it in the old way). In this way a new

generation can acquire of a set of underlying representations which differ from those which had been acquired by their parents. This might also mean that the children have one phonological process less or more than the older generation.

But it is equally clear that this is not the only possible locus of change in phonology, as among others, Lehmann (1973) and Kiparsky (1988) argue. There is no reason to believe that there can be no change in an adult speaker's phonology, and indeed the possibility of such change is widely recognised. Certain work in some frameworks has assumed that this can only involve the addition of a process, and in models where phonology is understood to involve a set of ordered rules, it has been proposed that such rule innovation can only occur 'at the end of the phonology'. As will become apparent in section 3.1, the phonological model that I adopt does not allow for lengthy derivations and so any case of process innovation will lead to processes which are of the same type as those which already exist in the phonology of a language.¹⁰

Change of this type, where a process is innovated into the synchronic phonology of speakers who already have a phonological steady state, could well have a greater potential for enlightening us about the nature of phonology, in fact, than changes which are innovated cross-generationally. This is because there is a clear potential for the phonological constraints that are active in the steady-state phonology of the speakers involved to influence the nature of the innovated processes. This can place restrictions on the nature of possible phonological innovations and can be taken to reveal aspects of phonological structure. This type of argumentation will prove important in some of the discussion in this thesis, particularly in chapter 3. The phonological changes that I deal with here are typically analysable as cases of innovation into phonological systems. I have little to say about change through acquisition. Other types of phonological innovation, apart from the introduction of an active phonological process, are also possible, of course, such as 'immediate' change to underlying segments, but I do not focus on this here.

The focus on this type of diachronic innovation of phonological processes provides a link between contemporary synchronic phonological data and historical data. This allows us to compare the two in a principled way, and legitimates the common approach mentioned at the

¹⁰ It will also be argued in section 3.2 that the status of phonology is very different to that of syntax, and this will be in line with the assumption that phonology is not necessarily fixed by a phonological Universal Grammar. This being the case, there is even more reason to believe that adult phonology can change than there would be if we took a strictly 'UG' approach to phonological acquisition. The fact that an individual's accents can change over time is further evidence for the idea that adult phonology can change.

start of this supersection (1.2). The evidence that we have for such past phonological processes is typically the result of the other main type of phonological change mentioned above. Phonological processes are often fossilised as diachronic changes in underlying segments; this occurs when a process is reanalysed by speakers and thus becomes lexicalised. Where the synchronic phonological process derived surface segments which differed from the underlying representation in only certain phonological environments this can lead to a segmental split, either to merge with an already existing segment or to create a new underlying segment (that is, in structuralist terminology, as a primary or secondary split, see Hoenigswald 1960). It will be important below that the phonological environment in which the ex-process occurred (its 'structural description') is also fossilised, as the 'exceptions' to a change. In this way the exceptions are expected to be explicable phonologically because they show where a previous phonological process did not occur. It is important, of course, through the use of philological evidence, to be sure that a process is accurately reconstructed and that possible subsequent changes are taken into consideration.¹¹ The diachronic data discussed in this thesis is largely comprised of cases where the lenition process involved has progressed from its introduction as a phonological process and has become lexicalised.

The data discussed in such 'long-distance' historical phonology, including much of that investigated in this thesis, often has an advantage for the theorist over synchronic data in that it is clearly categorical (in the sense of eg Lass 1997). It thus allows us to peer over sociolinguistic realisational variation in the search for phonological conditioning. This is, at least in part, a false picture, in that the introduction of a phonological process into a variety is typically accompanied by variation, as variationist sociolinguistics of the type reported on in, for example, Labov (1972) and Milroy (1987), has brilliantly shown. But in another sense, the picture is a true one. Just as theoretical phonology can abstract away from synchronic variation, so historical phonology can deal with phonological pressures and factors which can be seen to have been at work in the genesis and patterns of adoption of a change; long distance historical phonology, in fact, has no choice, as the past patterns of variation are inevitably lost.

¹¹ It is also important to be sure that grammaticalised relics of past phonology are not mistaken for synchronic processes. This means, for example, that 'initial mutations' such as those which occur in Celtic languages cannot be directly compared with truly phonological processes. These Celtic developments are further discussed in section 3.2.1.5.

Given the above, I place an important role in this thesis on the innovation of phonological processes by speakers and on the recognition of such processes in the phonological fossils which are left after a change in underlying representations. These, often with the aid of the comparative method of reconstruction and with the judicious interpretation of written evidence, can give us an insight into phonological processes of the past, which once, naturally, were part of a synchronic phonology. In section 1.3 I discuss certain further aspects relevant to the nature of the processes that I deal with. One key assumption is that they exhibit phonological ‘naturalness’ when they are introduced into a language, following an established tradition in phonology. I discuss the issues connected with this in the next section.

1.2.2.2 Naturalness, phonology and change

The notion of ‘naturalness’ is a good example of the way in which the concerns of historical and theoretical phonologists have interacted. It is a deep and difficult area to discuss, and some of the issues may seem not to be at the top of theoretical phonologists’ agenda, but it will prove important in this thesis, largely because of its historical bent, as I explain in this section. It will therefore be worth briefly considering the foundation of the ideas.

The concept of naturalness has been a fairly constant bugbear for theoretical phonology. The ‘natural’ reaction to the frequently perceived over-abstractness in Chomsky & Halle (1968) was one of the key motors in the development of phonological theory in the period immediately after its appearance (and the source of much overt debate, including those who spoke in favour of more naturalness in various ways, such as Kiparsky 1968 and Hooper 1976, and those who spoke up for abstractness, including Hyman 1970 and Bach & Harms 1972). Given the wide discussion of the issues that occurred decades ago and the diversion in the central course of phonological theory to other issues that has since occurred, it might not be thought worth mentioning at all. But I think some issues call for some treatment.

The imposition of naturalness on phonology can occur in more than one way. Either processes are expected to be phonetically natural, or the phonological entities that are assumed in a theory are expected to be phonetically natural. To an extent, these points are the same because the latter is a product of the former. The idea is that a process is more highly prized, (or thought to be more likely to represent psychological reality) if it accords with what is to be expected from the impact that might be had by the nature of the articulators or the acoustic properties of the speech signal.

There are several points which are worthy of short comment here: firstly, and most simply, we can note with Anderson (1985) that most phonologists do not think the issues surrounding ‘naturalness’ to be closed or convincingly solved. Writing in 1985, Anderson notes that “[r]ather what has happened seems to be that attention has simply been diverted by the exciting possibilities inherent in the major innovation of recent years: the enrichment of our notion of representation to include autosegmental and metrical structure” (1985, 350). It is ironic to consider these words from the current perspective of (especially American) phonology.

I see a fair proportion of the import of this thesis as a contribution to the understanding of representation and structure, which Anderson described as so exciting fifteen years ago. Needless to say, I do not consider these issues solved, but I am aware that other concerns are uppermost in other phonologists’ minds. The rise of Optimality Theory (the key texts are Prince & Smolensky 1993, McCarthy & Prince 1995 and see Kager 1999 and McMahon 2000b for different perspectives) has diverted the direction of phonological theory again. Segmental representations have a much less certain status in OT (there is some discussion of this here in section 3.1.2) but, as I show in this thesis, there is still much to be said in this regard.

Another point of interest here is the reawakening of interest in the issue of naturalness, which can also be seen to tie in with the rise of OT. Hale & Reiss (2000) resurrect some of the arguments for the view of phonology as an abstract computational system, which were first enunciated in the discussion of abstractness and naturalness in the 60s and 70s. They give them new and penetrative life in their critique of certain perspectives which have arisen in the Optimality Theory literature. The strangely not-quite-cyclic nature of this is intriguing. The ‘standard position’ in 1968 was quite abstract: phonology was basically conceived of as a calculatory system, and concerns of phonetic substance were not considered¹² and objecting voices were raised in favour of ‘naturalness’. Now the ‘standard position’ (in OT) is arguably ‘natural’, to the extent that phonetic substance is generally taken as the basis for the main type of entity in the theory (constraints are typically claimed to be ‘phonetically grounded’)¹³ and (minority) critical attention is directed against this.

¹² Except in the renowned chapter nine; see, for example, Stampe (1973).

¹³ It should be noted that other aspects of OT can become highly ‘abstract’ (eg, the use of Sympathy and Output-Output constraints, see McMahon 2000b), but it is the ‘natural’ part that Hale & Reiss (2000a,b) focus on.

I do not accept their final conclusion, but Hale & Reiss's (2000) argumentation is important. It is at least partly a reawakening of a point made in Bach & Harms (1972) and other older work which pleads for abstractness in phonology; one of their key points is that phonology should not be thought to be natural (ie, it should not be subject to the importation of phonetic substance) because the true locus for naturalness in phonology is in the innovation of phonological processes. It is here, they argue, that phonetic properties can affect the formation of processes and so phonetic 'substance' should not be imported into phonology, which should be allowed to calculate abstractly, and, over time, we might add, acquire 'crazy rules' which are entirely 'unnatural'. Their conclusion about the nature of synchronic phonology does not seem a necessary one to me but I defer discussion of this until section 3.1.1, where certain quite novel perspectives are brought to the issue. These will argue for accepting a generally 'natural' or at least 'simple' model of phonology.

I take an important point from the discussion here, however, in preparation for the preliminary points on lenition in section 1.4 and for the initial presentation of the Germanic data in chapter 2. This is Hale & Reiss's (2000) and Bach & Harms's (1972) point that newly innovated processes will be phonologically natural.¹⁴ There is an important caveat on this, though. The points made here about the naturalness of process innovation only apply to the processes which were innovated naturally. The tautology of this statement derives from the fact that phonological changes can occur in languages due to non-linguistic factors. In section 1.3, I discuss certain aspects of the distinction between the endogenous and exogenous innovation of phonological processes. Exogenous changes are those which will not necessarily show phonological conditioning on their introduction, and there is clearly no expectation that such changes will be natural because they do not arise from the phonological system into which they are imposed. It can be crucial to recognise this in any discussion of 'what is a possible process' because 'possible' can refer to the same thing as natural.¹⁵ This means that, when historical phonologists consider the empirical possibilities in the realms that they investigate and formulate hypotheses about 'what is a possible change', the understood caveat is always

¹⁴ It is probably not an unconnected point that some of the key theorists who argued for increased naturalness in synchronic phonology were historical phonologists (for example, Kiparsky; Dresher 1993 makes a similar point when reviewing Hogg 1992c). It seems that historical phonologists are more aware than most of the concerns of naturalness.

¹⁵ It can also involve certain other concerns, such as system naturalness, which is not based on articulation but on concerns of markedness and typology. This is relevant to the points taken from Hyman (1975) above because phonologists might think an endogenously innovated process which turns all occurrences of /p, t, k/ into /b, d, g/ to be impossible because it would level behind an unnatural system.

‘by linguistic, endogenous means.’ As may be expected, aspects of this will prove important in the final discussion of lenition processes; it has often been claimed that lenition processes are highly natural.

To sum up this section, concerns of naturalness play a key role in historical phonology and at least to an extent in synchronic phonology, too, because they are a major cause of the kind of processes that exist in synchronic phonology; they can be captured by the phonology (Kiparsky 1988 terms this ‘internalisation’). It is often the case that discussion of such issues focuses on articulation, and it seems likely that this is the chief influencing factor, but the notion of naturalness also allows for acoustic factors to play a role, as noted above. One clearly open issue here is whether such processes can become denaturalised once they are in the phonology of a language, and if so, to what degree. For the moment, I leave this question unanswered and simply note that there is an expectation that newly innovated endogenous processes will be natural. We return to these issues in section 3.1.

1.2.2.3 The quanta of phonological change

Many complex issues are raised by the consideration of naturalness in both synchronic and diachronic phonology. We cannot address them all here, but it will be helpful to briefly make one point explicit. This is the notion of the ‘quantum’ (as Lass 1997 terms it), ie, the unit for phonological change.¹⁶ The thinking behind this idea is straightforward, although the implementation is not always so clear: it does not seem possible to naturally innovate a phonological process which could change any individual phonological segment into any other.

There certainly seem to be constraints on how any segment can be changed, for example, we would not be surprised if a segment /e:/ became /e/ or /i:/, in one process, but we would be surprised if became /ʊ/ or /w/ or /tʃ/; the idea is that a change from /e:/ to /w/, for example, is not impossible in the history of a language, but it could not be achieved through just one innovation, ie, it involves more than one quantum. If we encountered such a correspondence between /e:/ and /w/ in two languages which we knew to be related, we would assume that a series of processes must have been involved, which had conspired to derive one segment from the other (or both from another proto-segment) and that these processes could all ‘move’ a

¹⁶ Trask (2000) describes the thinking discussed here as involving the proposal of ‘stepwise shifts’, and Picard (1999) discusses similar ideas under the heading ‘the minimality of phonological change’.

segment only one quantum; for example the set of processes might involve the following quanta: e: > o: > u: > u > w. The same naturally applies to consonantal change, and part of the discussion in chapter 3 involves a recognition of the quanta involved in certain lenition processes.

This idea is clearly related to the notion that a newly innovated process will be natural, the connection being that each quantum must represent a natural development. A key problem in this regard is that there is no clearly enunciated theory of quanta. As Lass (1997) comments, a theorists' understanding of what is a possible quantum will be related to their theory of subsegmental structure. As briefly mentioned above, I adopt in section 3.1 a model of phonology which views the building blocks of segments as privative units which are often 'larger' than the binary-valued distinctive features of Chomsky & Halle (1968) and much work which follows in that tradition. It might be contended that this supplies a theory of quanta, and to an extent it does, but these subsegmental units are largely motivated to capture articulatory information, and yet, as was mentioned in the last section, phonological change can be 'natural' in terms of acoustics as well as articulation. Acoustic factors can lead to changes in segments which are quite radical in terms of articulation; for example, we may assume that the change from /r/ to /ʁ/, which has occurred in a number of European languages, was almost certainly driven by acoustics. We will see below that the nature of the quanta of phonological change is subject to argumentation and investigation, as we would hope.

1.2.3 Terminology and concepts in historical phonology

The discussion so far has addressed the distinction in the different approaches to historical phonology taken in philology and in theoretical linguistics, and has begun to investigate some of the concerns of both of them, tending perhaps most towards the angle of theoretical historical phonology, a bias which will only partly be maintained in other parts of this thesis. Particularly in chapters 2 and 4, some of the concerns of philology will come to the fore. In the current section I return to consider aspects of the interaction between the two disciplines. This will focus in part on certain differences that have developed in the use of terminology, and also address some specific points which arise in connection with the symbols that are used for phonological transcription. The terminology used in the two disciplines does differ somewhat, and the two traditions have worked with quite different concepts at times, and these are not

always so simple to translate into the terminology of the other discipline. Partly because of the possible confusion that this could lead to in this thesis, I explain certain terminological points here and also take the opportunity to simply set out certain basic concepts of the approach that I adopt. The points covered in this section are unavoidably somewhat disparate.

One well known terminological distinction is Saussure's (1916) between 'diachrony' and 'synchrony'. We have already noted, though, that this is not the same as a distinction between 'historical' and 'present-day'. In places in this thesis, in an attempt to be specific, I refer to: 'diachronic phonological change' and 'synchronic phonological states' and in terms of the latter, I sometimes refer to 'past phonological states' and 'present-day phonological states'. This means that the term 'historical' can be used as a cover term for aspects of diachronic phonological change and of past phonological states; at times I make use of this helpful ambiguity.

Given that I explain in chapter 3 the model of phonology that I adopt in this thesis (with the aim of understanding both past and present-day phonological states), I do not deal here with all of the ideas which are explained there. A brief explanation of some of the fundamental points of the approach that I use in this chapter and the next may be in order, though. I assume an 'underlying' level and a 'surface' level to phonology, and hence also a set of phonological processes which map one onto the other. Both of these levels are comprised of phonological segments (which as we will see later, are comprised of 'elemental' subsegmental units). Despite the recognition of these levels, the segments on both are thought to be essentially the same type of entity ontologically. In terms of the processes in the phonology, I refer to segments as 'inputs' to these processes, or, if they have been affected by them, as 'outputs'

The philological and linguistic traditions differ slightly as to the symbols that they use to transcribe these phonological segments. In this thesis I explore some of the issues that lie behind this and the implications that it has for our understanding of the data. This is especially relevant for the data in chapter 2 which, as was explained in section 1.2.1, is taken from almost exclusively philological sources. I go on to analyse this data using the tools of theoretical linguistics, and this will involve a reinterpretation of it, but the initial presentation is faithful to the sources in terms of symbol use. This means that I make use of more than one set of symbols in this thesis, but the reasoning behind this move, and the steps taken to move from one set to the other are explained in detail during chapter 2; I hope that this is more enlightening than confusing. From chapter 3 onwards, I use the symbols of the International Phonetic Alphabet, except that I show the need to supplement these slightly in order to gain

phonological clarity of presentation once certain phonological distinctions (which are normally untranscribed) have been recognised, as is explained at the appropriate place (section 3.1.3.1). In the text, I normally refer to ‘stops’, not ‘plosives’ although nothing of import hangs on this.

In common with much work in the two traditions, I make a distinction between the symbols used to map segment onto segment in the formalisation of phonological generalisations. If the relationship between the two is intended to represent the input and output of a synchronic phonological process, I use an arrow with a shaft ‘→’ in informal rule-type formalisations of the processes involved (eg, $k \rightarrow x$). If the relationship is meant to represent a diachronic correspondence between two segments which is recognisable after the lexicalisation of a phonological process, I use the shaftless arrowhead ‘>’ (eg, $k > x$). I also use a third type of arrow where I mean the discussion to be relevant to both of these situations, this is ‘⇔’ (so $k \Leftrightarrow x$ means ‘a relationship exists between two segment-types which is due to a phonological process which either is synchronically active or was active but has now become lexicalised’). I naturally also make use of a distinction between slanty brackets for underlying forms (eg, /bul/), square brackets for surface forms (eg, [bulʲ]) and angle brackets for letters (eg, <bull>), although, where it is not necessary or might potentially be difficult to make a distinction, no brackets are used. At times, I also place IPA and other symbols between inverted commas, when I mean to refer to the symbol itself, rather than to the phonetic or phonological worth that it has (eg, to contrast the use of the symbol ‘p’ in philological texts for the same phonetic or phonological referent as the IPA symbol ‘θ’). I also use the colon to show that two segments (or series of segments) contrast in a language (for example, in English, /f, θ, s, ʃ/ : /v, ð, z, ʒ/).

Especially in chapter 2, I sometimes use traditional philological terminology to refer to series of stops. This comprises: ‘Tenues’ = /p, t, k/ (normally abbreviated to ‘Ts’); Tenues Aspiratae (‘TAs’) = /p^h, t^h, k^h/ (underlyingly aspirated versions of Ts); Mediae (‘Ms’) = /b, d, g/; Mediae Aspiratae (‘MAs’) = /b^h, d^h, g^h/ (underlyingly aspirated versions of Ms).

Later (in section 3.1.3.1), I change this convention so that the following holds: Tenues (‘Ts’) = phonological segments which are represented orthographically with the letters <p, t, k>; Mediae (‘Ms’) = phonological segments which are represented orthographically with the letters

⟨b, d, g⟩. The steps that will lead us to need these conventions are fully explained in the appropriate part of the thesis; the change comes about for good linguistic reasons.

A further terminological point which it may be well to address here concerns the term ‘aspiration’. This is probably most commonly used to describe the phonetic effect which is associated with the release phase of voiceless stops (in most phonological environments) in a wide range of languages including English, so for example it is the name given to the phenomena which is represented by the superscript [ʰ] diacritic in pronunciations such as [pʰm] *pin*. This is the way that I use the term in this thesis and the concept will be quite important. It is worth noting, however, that there is at least one other tradition of using the term to refer to something quite different. ‘Aspiration’ is also used to refer to what I call ‘debuccalisation’, that is, the loss of oral articulation in an obstruent,¹⁷ particularly in the case of fricatives. In Andalusian dialects of Spanish, for example, the process *s* → *h* can be observed (see, for example, Maza 2000a,b). This is often described as ‘aspiration’ in both philological and phonological literature, but I do not follow this tradition. The precise interpretation of the notion ‘aspiration in stops’ will form an important part of the discussion in section 3.2.3.1.

The conceptual difficulties that lie behind the contrast in the pair of terms ‘language’ and ‘dialect’ are well known, and I will not rehearse them in detail here (the terms ‘accent’ and ‘variety’ are probably just as problematic). It will be clear at several points in the discussion, though, that it is important to keep clearly in mind the status of the particular linguistic system that is being analysed. It is frequently not helpful in historical phonology to consider only ‘standard’ or ‘reference’ varieties as these are often at least partly ‘artificial’ and may have come into being through the mixture of ‘traditional dialect’ varieties, which means that they may not necessarily always have grown through natural phonological innovations (for the notion ‘traditional dialect’ see Wells 1982).¹⁸ Such reference accents are, for example, RP and General American for English (see Wells 1982) and *Standardlautung* for German (see

¹⁷ This will be an important type of process in several sections of this thesis and the details and implications of it will be investigated later (in chapters 2 and 3) so I will not discuss it in detail here. Probably coincidentally, we will see later that this process *is*, in fact, connected with ‘aspiration’ in stops, in a rather special way.

¹⁸ Essentially, Wells defines ‘traditional dialect’ as a “coherent alternative language variety” (1982, 4) and these naturally have a long history of being spoken in one geographical area where such linguistic innovations can accumulate to make the variety highly distinctive; ‘Northumbrian English’ is an example, as is ‘Bavarian’ (although these are obviously characterised by the existence of ‘varieties of traditional dialects’ themselves).

Mangold 1990), and it is clear that they do not have any special claim to phonological attention. Their phonology can also differ from that of non-reference accents just as much as it can from the phonology of varieties of other languages. At times in this thesis, we will consider points which are related to these issues, and we will be forced to consider the relationship between 'languages' and 'standard varieties of languages' and the historical connection between reference accents and non-reference accents. The interaction between varieties of a language and the development of a standard in spelling will also, at times, be relevant.

As a final point in this section, I consider one distinction which is frequently made in historical phonology, both by philologists and phonologists, between different types of phonological change. The same distinction is referred to in the literature by a number of terminological pairs, and I think this is at least in part due to the fact that a certain nicety in the distinction is not generally recognised. The distinction in question has been referred to as a contrast between 'spontaneous' and 'conditioned' changes, between 'combinatory' and 'non-combinatory' changes, between 'unconditioned' and 'conditioned' changes, between 'isolative' and 'conditioned' changes, between 'context-free' and 'context-sensitive' changes, and between 'paradigmatic' and 'syntagmatic' changes.

The same basic distinction is meant by all of these pairs. The first refers to a segmental change which was in some way independent of the phonological environment in which it occurred and the second to a change where the phonological environment played an important role. One problem here is that changes which are sometimes identified as belonging to the first group are not always entirely free from environmental concerns. For example, Jeffers & Lehisté (1979) classify the Germanic Consonant Shift as an unconditioned change, but quite a substantial part of this thesis will depend on the fact that phonological environment was not entirely irrelevant because there were clear 'exceptions' to at least some of the processes involved (this is dealt with in detail in section 2.1.1).

I think that this slightly problematic distinction can be better understood if we distinguish between three possible types of change in this regard. Basing myself on one of the pairs listed above, I thus refer in what follows to (i) 'conditioned changes', which can be thought in some way to be directly 'caused' by the environment in which they occur as a result of the influence of adjacent or closely neighbouring segments (clear examples are assimilations and dissimilations), (ii) 'weakly unconditioned changes', which are not entirely context-free in that they do not affect every occurrence of a segment in a language, but which cannot be clearly

seen to be ‘caused’ by the environment in which they occur,¹⁹ and (iii) ‘strongly unconditioned changes’, which affect every occurrence of a segment and where phonological environment clearly played no role at all. In what follows, I will occasionally use this three-way distinction where such clarity is needed. I turn now to the next section, where I continue to discuss quite foundational issues, but where I start to show precisely which are the particular aspects of historical phonology, of philology and linguistics, that I focus on in the rest of this thesis.

1.3 The focus and approach of this thesis

The field of historical phonology is wide, and as aspects of both philology and theoretical linguistics are relevant, discussion has centred on a large number of topics. In this section I discuss several of these and show how the concerns of the thesis relate to them. I spell out the focus that I adopt here and justify the fact that I set aside many of the concerns which are current in historical phonological debate.

Much of the work which was informally labelled ‘theoretical historical phonology’ above is not directed towards what Weinreich, Labov & Herzog (1968) label the ‘actuation problem’ in phonological change, ie, the factors which account for why certain types of change are innovated, but other types are not (along with the even more thorny question of why a specific change was innovated at a specific time and place). The key focus instead, is on the ways in which the phonological grammar changes in response to phonological innovation. King (1969) shows this when, in discussing a process which changed [hw] clusters to [w]²⁰ in many varieties of English (which he analyses as the addition of a phonological rule), he writes:

[w]e can, if we like, speculate on *why* this rule was added. Perhaps the speaker thought *w* sounded better than *hw*, perhaps *hw* was harder to pronounce than *w*. Such speculation is interesting but outside our immediate major concern, which is to give an account in our grammar of a change in speech habits.

King (1969, 80).

¹⁹ The word ‘cause’ here is in scare-quotes because the notions of causation and explanation are complex ones in historical phonology. There is some discussion of the issues involved in section 1.3.1.

²⁰ There is debate in the literature as to whether a cluster really was involved here, or whether the change was really [ɱ] to [w] (see, for example Wells 1982 on ‘Glide Cluster Reduction’ and compare the analysis in Giegerich 1992), but this does not alter the point made here.

The goal that King claims as his major concern is an important one, and it has provided for fruitful work in the tradition of theoretical historical phonology. It is this which has provided for certain connections with the aim that historical work can help in the determination of what is the best model of the grammar. But it is not a necessary restriction, however, and the key aim of this thesis is not that which King describes.

One criticism of the approach that King describes for historical phonology, and which has been pursued in much theoretical historical phonology relates to the status of the ‘actuation problem’, mentioned above, and concerns the ontological status of the type of account that is often given for such changes as [hw] > [w] into the grammar of languages. As many have noted, (for example, Hammarström 1978, Lass 1980, McMahon 2000b), the formalisation of these accounts as the introduction of a phonological rule, or by the reranking of phonological constraints is best seen as a (precise, informative) description, but not as an explanation of the process involved.²¹ To the extent that we can only claim true understanding of something if we can explain why it occurs, then the type of account that King (1969) advocates is only part of the story of the innovation of a process. Several traditions have recognised this and have worked towards providing accounts of phonological change which seek to provide some type of explanation for why a particular process was innovated and why others were not. These approaches try to go a little further than simply speculating that the output of a process might have ‘sounded better’ to speakers than the input.

Ohala (for example, 1974, 1992) has worked to provide accounts which are intended to be explanatory, attempting to root the innovation of certain phonological processes in measurable aspects of particularly acoustic phonetics. In doing this, he places an important caveat on such work: he focuses “only on sound changes that occur in similar form in languages distant from each other in geography, family membership, and typology” (Ohala 1992, 310). This is a good description of the kind of processes which are referred to as ‘lenitions’ and, because lenition is the centrepiece of this thesis, what I describe below is also consciously restricted to these types of change. We may also note that the description that Ohala gives for his restriction of focus is also, in fact, quite a good definition of the idea of ‘naturalness’ which was described above in section 1.2.2.2, and it is partly this which makes that concept relevant to this thesis.

²¹ There is work in this tradition which seeks to use potentially explanatory principles, such as the idea that rules tend to change order in a grammar to maximise feeding relationships (see, for example, Kiparsky 1982a) and that the principles of the Lexical Phonology model can help to understand the ‘progress’ of phonological processes through the grammar (see, for example, McMahon 2000a). I do not really engage with this work here, largely because of the ‘minimalist’ model of phonology which I adopt; this is addressed in section 3.1.

Ohala's approach is not the only one which seeks to explain why only certain processes are frequently innovated, whereas other types of process, which we could imagine should be theoretically possible, are not. This kind of reasoning was a key impetus for the development of enriched phonological representations of segmental structure, which started in phonological traditions in the 1970s (for discussion of the historical development of these approaches and some results, see, for example, Anderson 1985, Carr 1993, Ewen & van der Hulst 2001). These traditions include work in Feature Geometry (eg, Clements 1985, Sagey 1986) and in privative 'elemental' theories such as Dependency Phonology (Anderson & Jones 1974, Anderson & Ewen 1987) and Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990, Harris 1990). Some work in these traditions has attempted to provide a clear answer to the question 'what is a possible phonological process', and to the extent that phonological change can be seen as process innovation, this is the same question as 'what is a possible change?' As mentioned in section 1.2.2, much of the work in this thesis seeks to attach itself to these traditions. To return to King's (1969) example, the point is to have something to say about why [hw] might change to [w] and not something else, and to describe the mechanisms behind the processes involved.²²

We saw above that this approach is typically limited to accounting for 'natural' phonological processes. This implies, often implicitly, certain other restrictions of scope. When considering the introduction of a phonological change²³ into the speech habits of a community, it is common to distinguish in some way between the actual innovation of a process and its propagation in a speech community. In the terminology of Weinreich, Labov & Herzog (1968), this is a distinction between the 'actuation' of a process and its 'transmission'.²⁴ In this thesis, I deal exclusively with concerns that are relevant to the first of these ideas (the 'innovation', 'actuation', or 'primary change'); these are the phonetic and phonological factors which are inherent in the linguistic system and which provide the motivation and context for the phonological process concerned. The second idea in the terminological pairs (the

²² The deletion involved in the change *hw* > *w* might not seem too difficult to understand, but with other types of process, the mechanisms are not so straightforward, of course. In fact, if 'hw' was (or is) really a unitary segment /ʍ/, the process becomes more interesting here, as well.

²³ The same basic points are also taken to apply to other (non-phonological) types of linguistic change.

²⁴ The two concepts can probably also be equated with Sturtevant's (1917) distinction between 'primary change' and 'secondary change' and with Milroy's (1992) distinction between 'speaker innovation' and 'linguistic change'.

‘propagation’, ‘transmission’ or ‘secondary change’) is a vital companion to the first, without which widespread phonological change could never occur, but it is conceptually separable; it deals with the social factors which allow for certain innovations to ‘take root’ and spread through a speech community. Vital work in this regard includes Labov (1972, 2001) and represents an indispensable contribution to historical linguistics, but it is not inextricably connected with the phonetic and phonological factors which provide the innovations that can be picked up to mark group affiliation in a sociolinguistic manner and hence become spread through a community. To reiterate, I focus on innovation, not propagation, with the aim of explaining why certain specific innovations are possible and others are not.

A further restriction is necessary: I do not deal with one of the means through which processes can be innovated. There is another conceptual distinction to be made here which will allow a further tightening of focus for the thesis. It is common in work such as this to distinguish between ‘endogenous’ and ‘exogenous’ innovations (or between ‘internal’ and ‘external’ change). The discussion of the ‘factors which can motivate innovation’ in this section has been exclusively endogenous, and this focus is maintained throughout the great majority of this thesis. The key point of interest, as already explained, is on the enumeration of factors which are inherent possibilities in linguistic systems and which can lead to the innovation of particular types of process. It is certainly true that speakers can borrow aspects of the phonology of neighbouring varieties or languages, however, and the study of the effects of such types of contact has long played a role in diachronic linguistics. For example, Dixon (1997) reports that the phonological use of tone has spread among entirely unrelated languages in Sub-Saharan Africa as an areal phenomenon, and in a similar way, clicks are used in contiguous but genetically distinct languages in Southern Africa. These are cases where there is clear evidence that there has been exogenous phonological innovation in some of the languages concerned.²⁵

In certain parts of this thesis, I will appeal to the existences of exogenous innovation, but this is only where there is clear evidence that such influence could have occurred through contact. The main focus here is on endogeneity. At times, endogenous innovation has been rejected in favour of pure exogeneity, or the special case of exogeneity which Trask (2000) refers to as the ‘substrate explanation of change’ (where an innovation is accounted for as being due to contact with a language which was previously spoken in the area where the innovation is

²⁵ It is worth noting, of course, the use of tone and clicks was presumably innovated endogenously in at least one of the linguistic systems in these areas.

observed). As for example Trask (1997, 2000) points out, such ideas have been abused, and have been proposed as explanations for linguistic innovation, where there is no evidence that the supposedly exogenously innovated phenomenon was a feature of the supposed donor language. There clearly are such exogenous effects in diachronic phonology, as in other areas of historical linguistics, and where there is clear evidence for them then we can accept them happily, but they do not obviate the need to consider endogenous innovation.²⁶

One final distinction will help to mark out the focus of this thesis. The phonological processes that I deal with here are all ‘neogrammarian’ processes (which Labov 1994 also calls ‘regular sound change’) and were not obviously subject to ‘lexical diffusion’. This distinction has been discussed in some detail in historical phonology, especially in theoretical historical phonology, since work such as Wang (1969) and Wang & Cheng (1977) made a clear case for the idea that certain phonological processes proceed gradually through the lexicon, affecting, at least initially, one word or group of words, but not others with exactly the same phonological environment.²⁷ Such lexically diffusing processes go against the Neogrammarian’s ‘exceptionless hypothesis’ which was mentioned at the very start of this thesis.

In a detailed discussion of the topic, Labov (1994) reanalyses several cases of phonological innovations which had previously been thought to involve lexical diffusion, to show that they are in fact, perfectly regular, but with a fine phonological conditioning, which can be taken to indicate that lexical diffusion is less common than may be thought. He also concludes that there is no evidence that the types of process which are typically described as ‘lenition’ are ever subject to lexical diffusion. As we will see when we come to discuss the data in chapter 2 (and much of that in section 3.2.1), the processes involved there have long since been lexicalised into the underlying segments of the languages involved. In line with the discussion in section 1.2.2.1, this lexicalisation shows the signs of the phonological conditioning of the ex-processes, but it does not show effects of lexical conditioning; even if the processes were subject to lexical diffusion on their innovation, there is no sign of this. The null hypothesis in such cases is most likely that there was no lexical diffusion. In general, I do not deal with

²⁶ Even cases of similar, or identical changes in genetically related languages, are not necessarily a reflection of exogeny, even if the languages are spoken in relatively close proximity to each other. If the same phonetic and phonological factors influence similar languages, we should not be surprised that similar processes are innovated. It is this which accounts for ‘drift’, as discussed by Sapir (1921); see Nielsen (1981) for a discussion of some of the issues involved here with particular reference to the Germanic languages.

²⁷ For some discussion of the recognition of lexical diffusion and certain possible implications, see Kiparsky (1988, 1995), Labov (1981, 1994), McMahon (1994, 2000a).

variation in this thesis, neither with lexical variation, as in case of lexical diffusion, nor with the sociolinguistic variation which inevitably accompanies the introduction of processes.

In this section, I have put aside much of historical phonology, in order to be able to focus on a certain small but important, coherent aspect. This has involved the discussion of several key concepts in historical phonology. Much of the discussion in this thesis is directed towards an attempt to better understand the precise mechanics and patterning in the endogenous innovation of certain specific natural ‘neogrammarian’ types of phonological process (ie, ‘lenition processes’). I address questions concerned with what kind of phonological and phonetic factors lead to the innovation of these processes, but not others, and what, if anything, the group can be seen to have in common. To this end I focus almost entirely on processes which have been described as lenitions. I have yet to define what I mean by this term and I do not turn to this question until section 1.4, where I focus on the notion and present an initial definition. For the moment, we can define lenition in terms of what it is not: there is no discussion here of epenthesis or metathesis, nor either much of assimilation or dissimilation; the concept of segmental loss features briefly, as a possible case of lenition, and, in keeping with the ‘obstruent’ promise of the thesis’s title, there is no discussion of processes which affect vowels or sonorants.

At the start of this section, I referred to possible ‘explanations’ for the innovation as phonological processes and contrasted this with ‘description’. I believe that aspects of the coming discussion of endogenously innovated lenitions are explanatory, rather than simply descriptive. In order to see how they might be conceived of in this way, it is necessary to devote some attention to what is meant by the term ‘explanation’ in this respect. We turn to this in the next section.

1.3.1 Explanation in historical phonology

One of the key aims in this thesis is to explore some specific sets of diachronic data. These are notable and sometimes well-noted phonological changes which occurred in certain Germanic languages at certain specific times, and some of the characteristics that they show will be relatively clear from the preceding section. We can be quite sure that they occurred in some form or other (although the determination of the precise form that they took will be a fertile source for theorising in later chapters). We can make reasoned and reasonable conjectures as to when and where they occurred. I argue in this thesis that we can be aided in the exploration

of this data if we are armed with the insights of theoretical phonology. But we may well wonder whether these insights will allow us to truly *explain* the data.

In the last section, I briefly discussed the claim that is frequently made in the historical phonological literature that much historical phonology engages in description, and not explanation. This has been seen as problematic because it is quite widely proposed that we can only claim to have an understanding of a phenomenon once we have explained it. This problem has exercised historical linguists considerably, since at least Lass (1980)²⁸ and in fact it touches on a number of complex and wide-ranging issues, so many indeed that they cannot all be dealt with in one subsection here. The final defence of what I aim to do in this thesis is “it’s what other people do” in any case, which is not an unreasonable defence, given the notions of ‘paradigm’ and ‘research programme’ which are recognised in the literature on the philosophy of science (see for example Kuhn 1962). We cannot really hope to promote any progress in the philosophical understanding of the notion of explanation here, so I do not devote extensive discussion to the issue, but it will be salutary to briefly consider the background to the enterprise that I plan to undertake, else any claim to be engaged in ‘explanation’ will be baseless.

The basic import of Lass (1980) is that “explanations of sound changes in the strict sense do not exist” (1980, 42). Lass derives this claim from the assumption, which he himself has since described as “a bigoted, coarsely positivist assault” (Lass 1997, 332), that only one kind of explanation is real explanation - deductive-nomological explanation:

The ‘best’ explanation is ‘X, because it couldn’t have been otherwise (because Y)’ ... The explanation type that seems to come closest is the ... ‘Hempel-Oppenheim’ or ‘Deductive-Nomological’ schema, which characterises the physical sciences. It is based on deductive inference and, as its name implies, ‘laws’, and is ‘ideal’ in the sense that a well-formed explanation has the form of a deduction, and is in principle equivalent to a prediction.

Lass (1980, 9)

²⁸ Eaton & Koopman (1987) feature some historiographic background to the question of ‘explanation’ in historical linguistics and Lass (1980); and see Lass (1997) for his reflections of the issues. There are some famous quotations in the pre-Lass (1980) literature on ‘explanation in historical linguistics’ which are frequently given in discussions of the topic such as my discussion here: Bloomfield’s (1933) “the causes of sound change are unknown” (1933) and Postal’s (1968) “[i]t seems clear to the present writer that there is no more reason for languages to change than there is for automobiles to add fins one year and remove them the next, for jackets to have three buttons one year and two the next, etc.” (1968; 283). In terms of the possibility of explanation, these are generally reckoned to be a little pessimistic.

The key idea here is the connection between explanation and prediction. Lass (1980) maintains that, in order to be able to provide an explanation, for example, for a phonological innovation, we must be able to give a precise set of causes which are both necessary and sufficient to predict precisely when and where that innovation will unavoidably occur.

As many have pointed out in response to Lass (1980), however, the word 'explain' has more than one meaning²⁹ and the relationship between explanation and prediction is complex. One general conclusion is that deductive-nomological explanation is not the only type, and that other types are more relevant and useful notions for the discipline of historical phonology. Responses to Lass and more general contributions to the debate are found in, for example Vennemann (1983), Ohala (1987) and McMahon (1994, 2000b) and the papers in Koopman, van der Leek, Fischer & Eaton (1987); in one of these papers, Aitchison (1987) comments that

... the exact relationship between prediction and explanation is obscure, partly because of the vagueness of the terms 'predict' and 'explain'. Each of them is used to cover several different levels. Prediction can involve weak prediction (something is likely to happen), strong prediction (something will happen, though exactly when and where is unclear) and absolute prediction (something will happen, and the time and place can be specified in advance).

Aitchison (1987, 12)

Aitchison's points here are valid and are shared in part by Ohala (1987), who proposes that explanation in historical phonology should be deductive-probabilistic, ie, able to account for what is likely, not what is necessary. Intuitively, this seems right, because it seems unreasonable to expect that the discipline can predict exactly when and where any particular process might be innovated. Nonetheless, the kind of philosophical problems that Lass (1980) raised are worrisome. If we can only engage in Aitchison's 'weak predication', or even in 'strong prediction' as she formalises it, if we can only predict what is likely to happen, then the discipline has little claim to be viewed as scientific, or even as systematic, because there is no recourse if these predictions are not borne out. If we predict that something is only 'likely' to occur, then there is no possible cut-off point at which we can say that the prediction is confirmed or disproven. Philosophically, the prediction is vacuous, however correct it may seem, because it is not an empirical hypothesis.

There is a way of making Aitchison's (1987) 'strong predictions' empirically testable, however. We can maintain much of standard approaches to historical phonology but also

²⁹ As Stroop (1987) puts it, "[a]t times, it seems almost as if there are as many interpretations of the word 'explanation' as there are linguists" (1987, 259).

engage in real prediction and clearly approach an explanation of certain phonological innovations. The problem with Aitchison's notion of prediction is that it works the wrong way round. As Carr (2000) explains, basing himself on Popper's influential philosophy of science (eg, Popper 1959, 1963), a respectable definition of what makes a discipline into an explanatory science is that it makes falsifiable claims.

If we consider all three of Aitchison's understandings of 'predict', only 'absolute prediction' counts as falsifiable on her formulation. We have seen that a claim that something is 'likely' to happen ('weak prediction') is unfalsifiable, but also, even if we make the claim that something will definitely happen, but don't say when (as in 'strong prediction'), then if it does not happen during our lifetime, we can simply claim that this is not what we predicted because it could occur at any time in the future. If however, we make the types of prediction that Aitchison mentions negative, then 'strong prediction' becomes empirical, too.

Negative claims are much more easily falsified. If we claim that something will definitely not happen at a certain time and place ('absolute prediction'), that is clearly falsifiable. But also, if we claim that something will definitely not occur, but we do not say precisely when or where, then the thing in question can never occur, or the prediction is falsified. Just one counterexample means that the predictive hypothesis is at fault, unless, of course, the counterexample can be shown to be only an apparent counterexample and can be reinterpreted to be compatible with the prediction.

In this way, historical phonology can to an extent be explanatory. Lass (1980) seems to be claiming that only 'absolute prediction' is valid, and we can reject this, along with Ohala (1987), McMahon (1994) and others, while still acknowledging that what Lass (1980) discusses is important. While we need not necessarily agree with him that purely predictive deductive-nomological explanation is the appropriate comparative model for historical phonology, we can certainly agree with him that "[e]ven if one abandons strong deductive positivism, one still does not have to abandon certain criteria of logical and argumentative goodness, which essentially grow out of the same positivist tradition" (Lass 1997, 336). In what follows, as I hope will become apparent, I try to work with such traditions.

Certain predictions are made in this thesis (largely in section 3.2) in connection with lenition processes. These predict that, given a certain phonological background, certain types of change cannot be innovated by natural, endogenous means. These predictions are falsifiable, as I go on to show in chapter 5. These predictions are not perfect explanations of every aspect of what is possible in lenition, and what the exact patterning of every lenition process will be. Indeed we

cannot give such accounts. As McMahon (2000b) emphasises, *chance* clearly plays a role in process innovation. This is true both in terms of when and where particular processes are innovated, and also in terms of what the precise phonological conditioning of any process will be. We can predict, I show below, that if a lenition is innovated for any particular segment types, then it will *not* have certain characteristics (in terms of the segments and environments involved), ie, it could only be of a certain type. To that extent, we can explain why lenitions show certain types of patterning.

In other respects, in terms of a general understanding of other aspects of lenition (and many other aspects of historical phonology) we might agree with McMahon's (1994) conclusion, citing Bach (1974), on the issue of explanation: "we may have to accept a lower-key definition of explanation at a less elevated but more commonsense level: explanation might then constitute 'relief from puzzlement about some phenomenon'" (1994, 45). Here, we can bring together factors which correlate with each other in an attempt to understand as much as is possible about the processes.

In the next section, I turn to the notion of lenition for the first time in this thesis. The section features a general introduction to the notion and a quite brief description of what kind of phonological processes can be considered to be obstruent lenition. The approach adopted there will be far from explanatory, but it will illustrate certain previous approaches to the notions involved. The point of the section is principally to introduce the notions which will provide the background for the discussion of the data in chapter 2 and which will themselves be the basis of discussion in section 3.2.

1.4 A first look at lenition

Part of the point of the thesis is to undertake an examination of what historical and synchronic phonologists mean by the term 'lenition' and to consider whether the various ways that have been developed to understand the term are, in fact, insightful. In this connection, it will be important to consider whether the notion can be used to provide explanations of phonological events (in line with the discussion in section 1.3.1) and should be built into our model of phonology, or whether the notion is simply an invention of phonologists, which, while it might be a helpful metaphor, does not map on to anything that really exists in phonology, so that the phonological events will need to be explained with the help of other factors. Much of the

discussion of these topics will take place in section 3.2 and I will not pre-empt it here. This section is intended to provide an initial discussion of the term ‘lenition’, to attempt a simplistic, almost pre-theoretic account of what phonologists mean by it. This is to provide a backdrop for the presentation of the key sets of data in chapter 2 and will justify my inclusion of them as ‘candidate examples of lenition’. It will involve a brief discussion of the main points which are consistently mentioned in the existing literature on lenition, and will not come to any theoretically important conclusions.

It is now just over 100 years since the term ‘lenition’ was first suggested (in Thurneysen 1898)³⁰ and for that reason, if no other, it is reasonable to subject the term to a certain scrutiny, to ascertain what place it has taken up in the discourse of phonologists. As we will see, initially in the current section, and most notably in section 3.2.2, the term *lenition* is firmly established in the vocabulary of both diachronic and synchronic phonology and the meanings that have been attached to it by various writers have seen it play an important role in the development of several phonological theories. There has, however, been substantial disagreement as to what the term means, both in (simple) terms of which types of phonological process should be considered to be examples of ‘lenition’ and in (more complex) terms of what this tells us about the processes involved, that is, what they have in common that could lead us to describe them as being examples of one general process-type (we might say: as tokens of one type).

Lenition also goes by the name of ‘weakening’.³¹ The two terms are generally used interchangeably - most authors use them to refer to exactly the same thing and I shall do the same. Most authors who write on the topic choose one of the two terms and use it consistently; I will use *lenition*. When Thurneysen proposed the term, he was writing in German and he is explicit about the desirability, as he sees it, of using a Latinate loan-word (he also suggests the half-German *Lenierung*), but it is not surprising that a synonym from Germanic word-stock has established itself as a competitor in both English (*weakening*) and German (*Schwächung*). In the twentieth-century German-language texts which form much of

³⁰ Martinet (1952 and 1955) is scrupulous in ascribing the term to its inventor, and it is from Martinet that I learnt of Thurneysen’s creation of the term. Thurneysen’s suggestion comes in a review of Pedersen’s *Aspirationen i Irsk* (1897). Of course, it is by no means necessarily the case that Thurneysen discovered (or even invented) the *concept* simply by giving it a name.

³¹ As in Trask’s (1996) *Dictionary of Phonetics and Phonology*: “lenition /ləˈnɪʃn/ n. 1 (also weakening)” (1996, 201). Lavoie (2001) perceives a slight differentiation between the terms such that ‘lenition’ is more used in the context of historical phonology and ‘weakening’ more in synchronic work. This distinction would certainly fit with my usage here, but it is also the case in other work that the two terms are used synonymously.

the recent part of the tradition of philological Handbooks described in section 1.2.1 (which form much of the source material for the discussion in chapter 2), the terms *Schwächung* and *Lenition* are used interchangeably. In fact, both are often given, one almost as the definition of the other.

If we take Thurneysen's term in an etymologico-literal sense (he derives it overtly from Latin *lenire*), we would expect 'lenition' processes to 'soften', 'smooth', or 'calm' a segment (Kidd's 1957 definition of Latin *lenis*), so that the resulting segment is somehow smoother (for example) than the original. As we will see in this section and in later chapters, this etymological definition quite clearly fails to describe many of the processes which phonologists describe as lenitions (as such 'definitions' normally do). Given its Germanic etymological provenance, the term 'weakening' might seem more transparent in its meaning than 'lenition'. It is certainly true that notions of phonological weakness and strength are frequently tied up with lenition, but these, too, can be as unhelpful as 'smoothness' if they are not clearly explained and given an independent definition themselves.

The term is probably best known now in connection with (i) its use in the grammar of Celtic languages (as it was used by Thurneysen) and (ii) its use to describe some processes in the diachronic grammar of Romance languages. The term is also used much more widely, however. It is quite possible that Martinet's (1952, 1955) use of the term to compare the consonantal processes in these Romance languages to the ones described by Pedersen and Thurneysen for Celtic helped to ensure that 'lenition' was taken over from Celtic studies into wider use, but as it was already in use outside of Celticist circles by then (eg, in Paul 1944) we cannot be sure.

Whatever the etymology of the term, it is now commonly used in both historical and synchronic phonology. Either 'lenition' or 'weakening' (or both) feature in the index of a good number of recent phonological textbooks such as Hyman (1975), Lass (1984), Carr (1993), Kenstowicz (1994), Spencer (1996) and Davenport & Hannahs (1998) and in textbooks for historical linguistics such as Arlotto (1972), Hock (1986), Crowley (1992), McMahon (1994), Campbell (1998). There is also a large specialist literature on the topic. Several texts focus exclusively on the notion of lenition and its implications for phonological theory, and in yet others, it plays an important role. Some texts with a clear place in the lenition literature are: Lass & Anderson (1975), Lass (1976), Foley (1977), Escure (1977), Dressler (1985), Anderson & Ewen (1987), Bauer (1988), Donegan & Stampe (1989), Harris (1990, 1994), Elmedlaoui (1993), Kirchner (1998), Ségéral & Scheer (1999), Lavoie (2001). Along with

many others, these texts address the phenomenon from a wide range of perspectives. Considerations of lenition have been particularly fruitful in the Dependency Phonology and Government Phonology traditions, and it has been put to use quite frequently in the construction of phonological theories.

Given the widespread use of the term, it is a little surprising that there is substantial disagreement on how it is to be interpreted. Along with the idea that lenition has to do with changes in phonological weakness and strength, writers have connected it with the easing of articulatory effort, the loss of information in the phonetic signal, opening in oral stricture and with the loss of abstract phonological features. It is probably clear that not all of these ideas are compatible. In this thesis I seek to peel away some of the layers of meanings that have accumulated on the word ‘lenition’ and to come to a coherent understanding of what it really is.³²

One frequently discussed complication is the correlation between lenition and phonological environment. Lenition has also been linked to specific lenition-favouring environments which have even been labelled ‘lenition sites’. These are generally taken to be medial and final positions, or intervocalic and coda environments. It has been claimed that lenition only, or mostly, occurs in these environments. I return to these points in detail below.

It may be interesting to note how Thurneysen used the term when he introduced it. He proposed it to replace the use of ‘aspiration’ to describe the (morpho)phonological phenomena in Celtic languages which are now often labelled ‘mutations’ (although, especially for certain alternations in Irish, ‘lenition’ is still the term of choice, see, for example, Russell 1995). In all modern Celtic languages, there are synchronic alternations, which are fossilised, grammaticalised remnants of past processes of the type $p, t, k \Rightarrow f, \theta, x$, for example, (see Morris-Jones 1913, Ball & Müller 1992, Russell 1995; these are discussed further in section 3.2.1.5 below). Thurneysen suggests the term because he considers that all the processes involve a reduction in the intensity of articulation. It is also worth recognising that Thurneysen’s notion of lenition is not clearly tied to particular phonological environments. He suggests that we should speak of “a leniting final position” and of “a lenition of initial position” (Thurneysen 1898, 43).

³² It is interesting to note in this regard that lenition was the focus of the inaugural conference of the French national CNRS-funded Phonology Research Group (held at the University of Nice in 1999). One notable aspect of the conference was that, even after substantial detailed and high-level discussion, no consensus was reached on what lenition is.

1.4.1 Attempts to define lenition

It is possible to identify two main ways of defining ‘lenition’ in the literature. Of these, (i) sees lenition as a certain set of process-types, whereas (ii) claims that lenition is what happens in certain phonological environments. In terms of the vocabulary used to discuss the rules of Standard Generative Phonology, (i) can be seen as a classification in terms of the ‘structural change’ and (ii) as a classification according to ‘structural description’.

The foregoing discussion has already touched on both these ideas and it is true that many treatments of lenition in the literature combine aspects of both of them, but by no means all do. It is worth noting that the two definitions can be contradictory in terms of the kinds of processes that they include under the term, and also, (ii) is problematic because it runs the risk of becoming circular. Certain lenition theorists argue that a particular attested historical process need not be taken into account as an example of lenition because ‘lenition is what occurs in lenition environments and this example of what looks like lenition isn’t lenition because it doesn’t occur in lenition environments.’ The problem is that the notion ‘lenition environments’ is often defined as ‘those environments where lenition occurs’. Definitions of type (i) thus stand a better chance of initial success because at least they make a clear claim. For this reason, I take this tack for the remainder of this section.

One approach to (i) is to appeal to a connection with notions of segmental phonological strength. In this perspective, lenition is typically viewed as a move from a ‘stronger’ type of segment to a ‘weaker’ one. The idea is that relationships exist between classes of phonological segments which can be characterised in terms of their relative (segmental) strength. One frequently cited definition of such strength (in terms of its opposite, ‘weakness’) is Vennemann’s personal communication in Hyman (1975, 165)³³ “a segment X is said to be weaker than a segment Y if Y goes through an X stage on its way to zero.” This is basically a historical definition and it helps to tie in synchronic lenition processes with the types of process

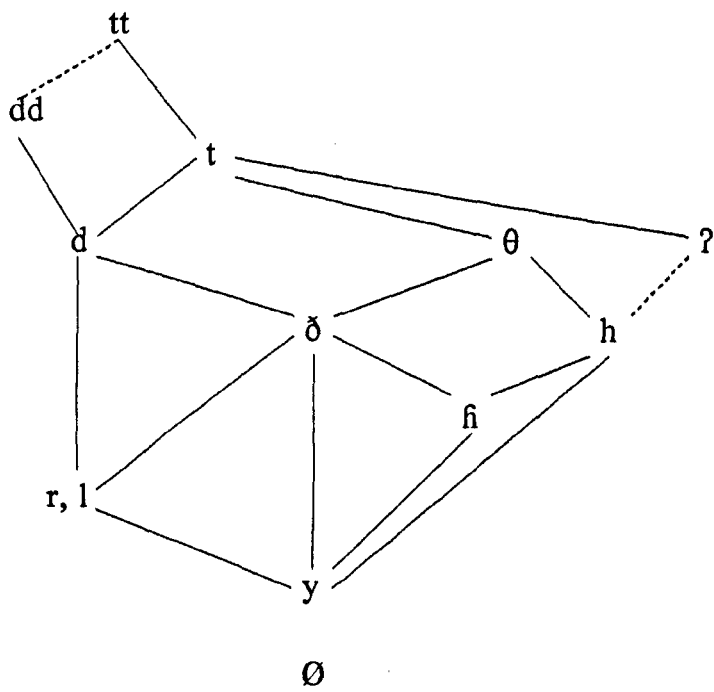
³³ This quotation must be a good candidate for the most cited personal communication in the history of linguistics. It crops up with extraordinary frequency in the literature on lenition (for example, in Escure 1977, Dressler 1985, Bauer 1988, Harris 1990, Harris-Northall 1990, Bloch-Rozmej 1995, Kirchner 1998, Ewen & van der Hulst 2001), a fact which we will return to in chapters 3 and 5. The original source is a remark by Theo Vennemann, as Hyman (1975) explains, and it has clearly struck a chord with phonologists. Hyman himself is well aware of the complicated conceptual issues that are associated with using this as a definition for lenition and he has since rejected it himself (personal communication). However, it is difficult to devise a better preliminary delimitation of the area of concern and this no doubt partly accounts for the frequency with which this short quotation is found as part of an initial attempt to define lenition in works which deal with the concept (including this one).

that are attested in phonological change: any process which has an input like Y and an output like X is seen as a lenition and the innovation of such a process is a diachronic change.

If we accept this definition, we can set up ‘lenition trajectories’ which tie in with the notion of segmental strength. This is the way that ‘lenition’ is used to group together a number of segmental processes because they are all perceived to have in common the fact that the output is weaker than the input. Processes of this type are then *lenition processes*. These types of processes are frequently innovated in diachronic phonology and they are typically described as being highly natural, which is arguably two ways of saying the same thing, as we saw in the discussion in section 1.2.2.2.

Quite a wide range of processes have, at one time or another, been described as lenition processes in this way. Hock (1986) has an extensive trajectory, reproduced here as (1.1) which includes a wide range of processes which he considers to be lenitions, illustrated at the dental place of articulation. Any movement downward is considered lenition and dotted lines indicate processes which he considers to be possible, but not observed:

(1.1)



But this is overwhelming. As Kirchner (1998) argues, it is not clear that all of these processes are in fact possible and, in any case, the trajectory deals with a range of segments which fall outside the obstruent-based focus of this thesis. It does provide a starting point, however, and a range of candidate processes that we can consider, although it also excludes

processes that others include. One important point that we can take from Hock's (1986) chart, however, is that lenition is not generally thought to change the oral place of articulation of a segment. Lenition processes have been argued to alter the length, voicing or manner of articulation of a segment, but not its place. I take on this aspect of Hock's definition and subsequent discussion will exclude any consideration of processes which majorly change the place of a segment.³⁴ This point is not vitiated by the inclusion of debuccalisation processes such as $\theta \Rightarrow h$ on Hock's trajectory, nor in the coming discussion, because debuccalisation is perceived to involve the removal of a segment's oral place of articulation, not its change. Despite Hock's exemplification at one place of articulation, it is generally recognised that lenition processes often affect a whole classes of segments, for example, all the stops in a system, or all the fricatives. Hock's trajectory also includes a final stage to \emptyset (ie, deletion). This is frequently taken to be the ultimate loss of segmental strength.

In common with most discussions of the concept, in this thesis I do not discuss every single type of process which has at one point or another been described as a 'lenition' because constraints of space preclude it. For example, some authors have proposed that vowels can undergo lenition (eg, Dressler 1985, Bauer 1988) but I shall not deal with that idea here. There are good reasons to believe that vowels and consonants behave very differently in this regard (not least the fact that the two types of segment occur in very different phonological positions) and, in keeping with the title of the thesis, I focus on lenitions which affect and produce obstruents. The sets of data to be presented in detail in chapter 2 can all be, and at various times have all been, described as examples of obstruent lenition processes.

In the lenition literature, several attempts have been made to derive the notion of phonological 'strength' from other phonological concepts. These include a segment's inherent sonority, its patterning of sequencing in syllabic constituents, segmental complexity, the 'openness' or degree of resistance in the vocal tract that is associated with the segment and the notion of a segment's perceptual salience. I investigate some of these notions in section 3.2.2

³⁴ Some writers have sought to extend notions of phonological strength to place. Escure (1977) and Foley (1977), for example, have constructed strength scales which refer to this parameter, but many other writers on the topic have taken issue with their suggestions (eg, Cohen 1971, Smith 1981, Harris 1985, Anderson & Ewen 1987) and have comprehensively demolished them. Honeybone (1995) is a largely failed attempt to address some of the issues brought up in the discussion of 'lenition of place', but it does include some argumentation to the effect that place is not like manner or voicing in terms of lenition. The hedge 'majorly' will be explained in due course; it relates to the notion of major articulators.

where I return, in quite some detail, to the thorny question of how to provide a coherent definition of something that we might want to call lenition. I argue there for a novel position which addresses the issues from a less monolithic perspective, but we need a preliminary definition of the concept which tallies with the general consensus and which will allow us to turn to some data.³⁵

The definition of lenition provided here will simply involve a set of lenition trajectories, constructed partly in accordance with Vennemann’s dictum on strength and weakness. These trajectories rely on common historical weakening paths (and are in fact abstracted away from Lass & Anderson 1975, which is a classical locus of the discussion of lenition), to feature just those processes which involve clearly simplex obstruents. Lass & Anderson (1975) recognise two basic types of lenition processes: ‘opening’ processes and ‘sonorization’ processes. Opening involves “progressive continuantization without a change in glottal attitude” (1975, 159) and sonorization includes voicing. These are shown in (1.2a). From them, we can construct segmental lenition trajectories such as those in (1.2b). These put together the types of lenition and include the idea that “[t]he last stage in any lenition is deletion” (Lass & Anderson 1975, 159).

(1.2) a) Types of obstruent lenition:

Opening: stop \Rightarrow affricate \Rightarrow fricative \Rightarrow glottal \Rightarrow deletion

Sonorization: voiceless \Rightarrow voiced

b) Lenition trajectories exemplified at the velar place of articulation:

$k \Rightarrow kx \Rightarrow x \Rightarrow h \Rightarrow \emptyset$

$k \Rightarrow g \Rightarrow \gamma \Rightarrow \emptyset$

With trajectories like these, the set of lenition processes can be summarised and exemplified as in (1.3), although it should be remembered that lenition processes often affect whole classes of segments:

(1.3) Lenition processes

affrication: $t \Rightarrow ts$, $k \Rightarrow kx$

spirantisation: $k \Rightarrow x$, $d \Rightarrow \delta$

³⁵ I also develop (in section 3.2.3) a different notion of ‘strength’ which is quite unlike the way the term has been used in this section, but that will be clear in its context.

voicing: $f \Rightarrow v$, $p \Rightarrow b$

debuccalisation: $s \Rightarrow h$, $x \Rightarrow h$

Most of these are uncontroversial. Some authors have rejected affrication as a case of lenition, but this seems wrong to me and, following the Anderson/Lass/Ewen tradition, I include it.³⁶

At times, writers on the topic of lenition also discuss ‘fortition’, which is seen as the ‘opposite’ of lenition. It is claimed that any process which moves a segment up a lenition trajectory, rather than down it, is a fortition, which would mean that fortitions would have the opposite effect to lenitions (for example $ts \Rightarrow t$, $h \Rightarrow x$, $x \Rightarrow k$ would be fortitions). However, the attested occurrences of such processes are extremely few. There are generally so few, in fact, that we might doubt whether they can be naturally, endogenously innovated at all. For this reason, I do not discuss the notion that there might be such a thing as fortition in any detail in this thesis. We shall see below that certain cases of ‘fortition’ which have been reported in the literature are not fortitions at all, and it seems likely that other putative fortitions can also be better analysed in different ways. Lenition processes are extremely common, and it is they that demand our attention.

1.4.2 The place of lenition in this thesis

In the chapters to follow, I investigate cases of attested phonological processes which fit onto both of the types of lenition trajectory just discussed: ‘opening’ and ‘sonorisation’. I also revisit the whole idea of lenition in an attempt to discover what lies behind lenition processes and to what extent we are justified in viewing it as a unitary phenomenon. The discussion in this section will suffice to delimit the area of interest for the moment, and we can now move on to discuss some real examples. As has been mentioned already, these are taken from the diachrony of certain (chiefly West) Germanic languages and often involve well-known cases of lenitions. Chapter 2 investigates the history of these processes in quite some philological detail because it is important that the data used in later theorising are secure. Chapter 3 subjects the general notion of lenition to scrutiny, and we will see that some of the formulae used in this section, such as ‘ $k \Rightarrow x$ ’ and ‘ $t \Rightarrow d$ ’, are not as simple as they initially seem.

³⁶ The justification for this step, if it is needed, will become apparent in section 3.3.2.

2 Lenition in Germanic languages

The material in this chapter is chiefly expository. Some key sets of data (which we will return to in later chapters) are presented here in a way which will be quite far removed from the final form that they will take in this thesis. To a large extent, the presentation in this chapter is of the type found in Germanic philological Handbooks and Grammars (as defined in section 1.2.1). This chapter is, then, the main locus of 'traditional' philological discussion in the thesis; it is the only place where external historical concerns are considered, such as the dating and geography of particular phonological changes.

The data which I discuss are hardly new and it is not intended that novel evidence from manuscript or comparative sources is to be considered. In the coming exposition, I thus rely on the results of the traditional philological research rather than on any personally gathered data. For each set of data, I give some brief comparative evidence, taken from the Handbooks and other secondary sources, but this is largely for the sake of completeness. The philological tradition has established beyond doubt that, for each of the cases considered, there is some diachronic phonological event to be accounted for.

This is not to say that the way the data is presented in any of the Handbooks consulted is to be taken as sacrosanct. Every researcher imposes their own perspective on the data that they consider and, in any case, there is sometimes philological disagreement about the data discussed in this chapter. I illustrate this in places, and show how different interpretations of the philological facts could have an influence on the discussion of the data in future chapters. The discussion of such points in this chapter is sometimes necessary in order to establish that the changes which are discussed here really did occur and that the data sets are a good basis for theoretical discussion.

There must be some explanation for the comparative data which is given in this chapter to back up the phonological processes discussed here. This data includes cases when writers have started to spell words in a way that is different to what was conventional before. It also includes cases where a group of words are spelt in one way in one language and are spelt in a different way in a language which we take to be genetically related. For some of the processes described here, it involves the comparison of present-day non-standard dialect forms which differ from each other or from the standard form. The normal explanation, which is accepted in this chapter, is that such situations represent the fact that phonological processes have been innovated and that these processes have changed the phonology of the languages in question,

often leaving their mark as fossilised, lexicalised changes in underlying segments. In later chapters, these explanations are questioned for some of the processes described here.

In the exposition here, I settle on what I take to be a distillation of (at least one of the) traditional ‘received positions’ on the data from a philological, Handbook-style perspective. As the reader encounters these presentations, certain aspects of them will likely appear out-dated.¹ Some of the problems with the analyses presented here are well-known and will be addressed later. My discussion of such points is by no means intended to imply that I consider the recognition of them to be an achievement on my part, but there will be good reason to discuss them in this thesis. Equally, the decision to present highly traditional versions of the data discussed in this chapter was taken advisedly, as will become apparent in chapter 3 and elsewhere.

It is philologists who have considered the data in most detail in their attempts to write the phonological history of the languages concerned, and, in subjecting it to such scrutiny, it is they who are likely to be most faithful to what they perceive to be the objective reality of the changes. In this chapter, I follow them respectfully. When we return to the data in later chapters, novelty begins to arise. Some of the general theoretical conclusions which will be drawn about the type of processes which they exemplify are largely compatible with the data as presented here; some of the coming reanalyses of these data are reinterpretations, rather than radical revisions of history. Some of the linguistic evidence which is brought to bear on the data, however, will require a clear re-writing of the linguistic history given in this chapter.

The data is presented in four main sections in this chapter (sections 2.1.1 - 2.1.4). Each of these sections works through a set of philological data of the type described above. The data discussed in these sections includes the key lenitions in this thesis, although, as we will see, the wider data sets sometimes include processes which are not easily interpretable as lenitions. Much of the contribution to knowledge of this thesis is intended to derive from (i) the implications of aspects of phonological theory for these lenitions and (ii) the implications that these lenitions (and others) have for phonological theory. These two points will lead to (i) the reanalysis of certain aspects of the lenitions and (ii) certain novel proposals about what is

¹ The adoption of a philological approach in this chapter will not mean that the exposition is atomistic. It is true that many of the Handbooks and Grammars of the tradition in question here (eg, Braune 1891, Paul 1916, Kienle 1960) devote many pages to the history of individual segments, but this is also preceded by a section on ‘general developments’ where the author discusses general processes which affected classes of segments.

possible in phonology (and in historical phonology in particular). The four sets of diachronic events described here are thus discussed both as changes in their own right and as attested examples of general types of change. The lenitions featured here are compared to analogous processes in other languages in later chapters (especially in section 3.2.1).

This chapter also includes a short summary section after the presentation of the data sets (section 2.2) and then follows an important section (2.3), which starts the process of phonological interpretation for the data. This partly involves confronting the data with quite simple aspects of a universalist phonological approach (which will affect the symbols used to discuss it), and partly involves reconsidering aspects of the data in line with notions which focus on the need to ensure the ‘reality’ or ‘phonological plausibility’ of the reconstructions that we assume for past phonological states in line with a strict application of the uniformitarian principle.

2.1.1 Germanic obstruent lenitions

This section discusses four diachronic phenomena. These range from what is probably the best known set of changes that philology has ever recognised to some data sets which are much less well-known. Two of the changes discussed here have affected the varieties of languages which have come to form the basis of present-day standard languages and the other two have occurred in varieties which have not (both of these are still attested in living non-standard varieties, however). The effects of one of the sets of processes can be seen in all Germanic languages, the effects of the second can be seen in the standard present-day German language (ie, High German dialects of West Germanic), effects of the third can be seen in southern varieties of English and of the last in a range of central and southerly non-standard dialects of German.

These processes are known by a range of names. I refer to them here as: the Germanic Consonant Shift, the High German Consonant Shift, English Initial Fricative Voicing and Inner-German Consonant Weakening.² It is important to recognise from the outset that the four sets of data discussed here are by no means necessarily to be conceived of as four coherent

² The other names by which they are known and the implications and reasoning behind the choice of these designations are discussed in the individual subsections below.

processes. It is conventional to treat a set of events together under each of these four headings, and I follow this tradition here. As we will see, however, the phenomena collected under some of these labels are quite heterogeneous and are largely grouped together by tradition, or by the fact that they occurred at approximately the same stage in the development of a language.

The four sets of changes can be seen as connected because the Germanic Consonant Shift (henceforth often 'the GCS') created segments which became the input to the other three processes (which will from now on generally be referred to as the 'HGCS', the 'EIFV' and the 'IGCW'). The GCS is an entirely reconstructed set of processes (a product of the comparative method) in that we have no written records of either the language which formed the 'input' to the change (Indo-European) or of the language which was the result (Proto-Germanic). The other three are of a different status because we have written records which, to a greater or lesser extent, document the changes in question. At least the output of the changes (that is, the novel phonological segments which were created through the introduction of the processes) are recorded in texts.³

It is worth exploring, at least briefly, the rationale behind the selection of specifically these four processes for special consideration. The chief reason is that they exemplify some processes of the types which were identified as 'lenition processes' in section 1.4. In terms of the taxonomy of process-types discussed there, the four processes investigated here are typically described as featuring voicing, spirantisation, affrication and debuccalisation (the question of geminacy also briefly arises). The four data sets featured here are well-researched examples of diachronic phonological events which can be used to investigate the kinds of processes which have been labelled as lenition, although, as we will see, they are not necessarily the most 'obvious' cases of lenition. I see this as a further reason to examine them in this light, for we can probably learn little new about what we call 'lenition' unless we confront it with complicated cases. The fact that they are all connected, in that the output of one is the input to the other three, is another important reason for examining precisely these four sets of data. It is thus in fact important to treat them together, so that their true context can be understood.

³ This means that there are manuscripts where words are spelt in ways which are most easily interpreted as evidence for the fact that the changes in question must have occurred in the phonology of the author of the texts. It is assumed that this influenced the author (either consciously or subconsciously) to write with an orthography which differs from the orthography which would better represent the phonological system which is assumed for a previous phonological state.

It is also notable that much of the data discussed in this chapter is ‘classic’ data. By this, I mean that it is well-known to phonologists (especially Germanicist historical phonologists); indeed, aspects of it have been discussed extensively for centuries, practically back to the foundation of systematic historical phonology in the early 19th century. The fact that this thesis works with such ‘classic’ data need hardly be seen as a problem, however. It would only be problematic if the data’s iconic status meant that an analyst were tempted to treat it with too much respect, either in terms of accepting previous analyses or, indeed, in terms of blindly accepting the data as it was described centuries ago. A danger does exist that classic data could become set in stone. Were this to occur, it would not only put a halt to any development in the understanding of the data, it might also have the effect that the data and classic analyses of it come to be viewed as simply being ‘true’ (rather than having the status of a hypothesis). It could then be mistaken as a rock-solid foundation for theory building such that a theorist does not examine these foundations for possible cracks. One of the aims of the thesis is to reproblematised this classic data. This is done on the assumption that such data, such as the GCS or the HGCS, cannot be considered only in its own right (as can be the case in a philological approach), but must be compared with what we know about other languages and language in general. My aim here is nothing new in itself, of course, but it plays an important role in the consideration of data such as some of that discussed here, which has so iconic a status.

A further result of the fact that some of the data discussed here is ‘classic’ data is that there has been multitude of analyses of it. It would be an almost super-human task to consider everything that has been written on some of these processes. Luckily, others have already done this and presented the results in an easily accessible form (for example, Schrodts 1976 and Collinge 1985). I discuss certain previous analyses of these processes here (in this chapter and also a little in chapter 4), focusing partly on those which present a particular challenge to the traditional view, or are notably mistaken or insightful. Needless to say, despite the vast literature on some of these processes, I do not believe that the correct analysis of them has yet been proposed. The data discussed here will be comprehensively re-interpreted in chapter 4; the ‘full story’ will thus emerge as a synthesis of the discussion in the current chapter 2 and coming chapter 4.

As was briefly mentioned in section 1.2.3, the phonological symbols used in this chapter are not consistently those of the IPA (or another linguistic) tradition. Rather, I use the symbols that are conventionally found in the philological Handbooks and Grammars which form my

main source. These symbols often do not differ greatly from their IPA equivalents and most of those that do should nonetheless be relatively straightforward to interpret for the phonologist who is not familiar with the philological literature (eg, ‘b, d, g’ are voiced fricatives and ‘p’ is the voiceless dental fricative). As we will see in section 2.3, the symbols used in such description can be misleading. In fact, however, I will argue in section 3.1.3.1 that this can also be more generally the case for the symbols used in the contemporary synchronic phonological literature. As this point will form a major aspect of later discussion, I will not preempt it here, but it should be borne in mind that the description of the changes given in this chapter is to be substantially revised. Because the discussion in later chapters is more easily described as ‘historical linguistics’ rather than ‘philology’, one of the first acts of reinterpretation for the four data sets will be to replace these symbols, where necessary, with more IPA-like ones. It may thus seem a little odd to use non-IPA symbols here. However, the issues are not all as straightforward as they may appear and there are certain issues which the use of philological symbols clearly bring out. These are addressed in detail in section 2.3.

2.1.1 The Germanic Consonant Shift

The set of processes which I refer to here as the ‘Germanic Consonant Shift’ are one of the defining features of the Germanic grouping of languages. This is because they are typically thought to have been innovated in the variety of Indo-European which forms the direct ancestor of all Germanic languages on family-tree models of linguistic relationship (and which is sometimes referred to as ‘Germanic’, so that the term refers both to a linguistic grouping and a language which was once a synchronic reality, although the term ‘Proto-Germanic’ is often used for that latter). Once the segments which were derived by the process were lexicalised, there were clear and distinct correspondences between underlying segments in many words of Germanic and etymologically related reflexes in other IE daughter languages. Naturally it was these which first allowed scholars to recognise that the changes had occurred. Some of these corresponding segments are still highly visible in modern Germanic languages (but much less so in the ‘High’ dialects of German where later processes have substantially changed more of the outputs of the GCS than in other Germanic dialects, as we will see in the next section).

In fact, the GCS is probably the single best known set of data in historical phonology, although it is also often given a different name, ‘Grimm’s Law’ (for example, by von Raumer

1837, Collinge 1985, Iverson & Salmons 1995, Labov 2001), after one of the first scholars to recognise and describe the processes.⁴ My choice of title for the data is not intended to have much theoretical import and is in any case, probably the most commonly used name in the German philological literature. It is true that ‘Grimm’s Law’ would do as well as ‘the Germanic Consonant Shift’, because ‘law’ is probably as semantically bleached as ‘shift’ (I simply use the term ‘shift’ as a relatively vague designation for ‘a substantial phonological change which involved several segments’). I would not want to ignore Grimm’s extraordinary contribution to the discipline in which I work, but, if reason be needed to justify my choice of name, it seems reasonable that the name of a process should focus on the phonology rather than on a philologist. It might also be noted that I come to very different conclusions in the final analysis of the data discussed here to those assumed by Grimm.

2.1.1.1 The GCS: take 1

The GCS can be conceived of relatively uncontroversially as a set of several separate innovations (I assume this here, although some previous analyses have insisted that there is an underlying unity to the developments; see the discussions in van Coetsem 1970, Schrodtt 1976, Collinge 1985). The introduction of the GCS phonological processes altered the realisation of practically all the Indo-European obstruents in the varieties of IE into which it was innovated. There were few fricatives in IE and the vast majority of the stops were affected (most occurrences of the stop segments changed, although some environments prevented the processes from occurring, as we will see imminently). The oldest, ‘classical’ reconstructions of

⁴ We have already seen (in section 1.2.2) an example of the fact that this data is used very frequently even in texts on general phonological theory (eg, Hyman 1975 - it is also mentioned, for example, in Sommerstein 1977, Lass 1984 and Kenstowicz 1994). One probably non-hyperbolic claim regarding the data is that “[i]f non-specialists know anything about historical linguistics, it is Grimm’s Law” (Lehmann 1967, 46, quoted in Collinge 1985). It is well recognised, though, that Jacob Grimm was not the first to notice or describe the correspondences; indeed it is difficult to be sure who was. It is normal to mention Rask’s (1818) discussion of the topic, but Lehmann (1993, 290) points out that Schlegel (1806) also discusses some of the simple correspondences. In any case, it seems that the term ‘Grimm’s Law’ is more common in writing on the topic in English than the German equivalent (*das Grimmsche Gesetz*) is in discussion in German. Another option (especially in German) is to call it the ‘First Sound Shift’ (*die erste Lautverschiebung*) which derives from Grimm’s own writing (as, for example, Vennemann 1984 explains); this is intended to contrast with the ‘Second Sound Shift’ in the history of the German language, which, as we will see in the next section (2.1.2), only affected certain dialects of German. It seems a little germanocentric to use these names, so I shall persist with ‘Germanic Consonant Shift’.

IE stop consonantism, which is that most simply derived from the comparative method is shown in (2.1). It has four series of stops and five places of articulation (or, rather, four simplex places and one double articulation). This is the inventory of Brugmann & Delbrück (1886-1900) and is commonly assumed in the standard IE Handbooks⁵ (see Lehmann 1993).

(2.1)

Tenues Aspiratae	ph	th	kh	kh	kh ^w
Tenues	p	t	k	k	k ^w
Mediae Aspiratae	bh	dh	gh	gh	gh ^w
Mediae	b	d	g	g	g ^w

The names given to the four series were briefly discussed in section 1.2.3 and it will prove convenient to use them here (often abbreviated to ‘*TAs*’, ‘*Ts*’, ‘*MA*s’ and ‘*M*s’) to refer to the proposed class of segments, in part, following Schrodts (1976) compendious discussion of research on the GCS. The segments which most clearly demand further attention are the *TAs* and *MA*s and they will receive it in section 2.3.2. It is clear that, in some sense, the four-way symbolic distinction at all places of articulation is intended to represent a four-way laryngeal contrast. The symbols used here, with non-superscript aitches are common in philological work.

The symbols / kh, k, gh, g / are used to represent palatals (of the type which would be represented by IPA / c, j /). In the Handbooks these are assumed for common Indo-European on the evidence of attested segments in the so-called ‘satem’ group of IE daughter languages, such as Indo-Iranian, Slavic, Armenian, Lithuanian. Given that Germanic does not belong to this group, we can exclude this class of segments from further consideration here, as is quite common in Germanic Handbooks (eg Luick 1914-1940, Krahe 1969, van Coetsem 1970). The Handbooks generally exclude the palatals from discussion of the GCS because it is claimed that

⁵ It is not so commonly assumed in the Germanic Grammars, however, for reasons which will be explored in section 2.3; Luick (1914-1940) has exactly this system, however, as does Kienle (1960), Ramat (1981) and Bammesberger (1992).

the contrast between them and the velars was lost in pre-Germanic, before the GCS occurred, so that the two classes merged completely to become velars.⁶

The segments / kh^w, k^w, gh^w, g^w / are 'labiovelars' (or 'labial-velars'). These are typically assumed in the Handbooks for pre-Germanic IE (although, as many explain, eg Luick 1914-1940, Prokosch 1939, they typically lost their labiality, to become plain velars before /u, o/ and consonants and lost their velarity before /i, e/). These segments did serve as the input to the GCS processes and labialised outputs are attested orthographically in Gothic (see, for example, Braune & Ebbinghaus 1981). While various processes thus reduced the number of truly labiovelar reflexes of the IE labiovelar stops, I nonetheless include them in discussion here without further comment. Nothing in their status alters any of the points to be made below.

If we assume four series of IE stops as the input to the GCS, the four are reduced to three contrasting series after the GCS. The segments which were the result of the GCS changes are the same for the *TAs* and the *Ts* (another way of expressing this is to say that the *TAs* and the *Ts* merged in the GCS). The simple segmental changes can be represented as in (2.2), where the sameness of the output for IE *TAs* and *Ts* is recognised by labelling them GCS_{1a} and GCS_{1b} (so they can be referred to together as GCS₁); it should be remembered that, even though all the changes are grouped together here expositively, they are to be conceived of as a set of processes which, while similar in their extent and time of occurrence, are not necessarily linked. The order here is not intended to make any claim as to their relative chronology:

- (2.2)
- GCS_{1a}
 - ph > f
 - th > þ
 - kh > χ
 - kh^w > χ^w

⁶ It is a little inconsistent to exclude the palatals while continuing to consider the *TAs* (as where these are considered, they are claimed to have merged with the *Ts*) but nothing substantial hinges on this decision and it is done for good expository reason. In fact, it is highly likely that the palatals were never a feature of common IE. It is very uncommon for a language to have three contrasts involving the dorsal articulator (Maddieson 1984 counts only 4 languages out of the 317 in the UPSID database which have a velar stop and a further 'body place' stop and a labial-velar stop). It is entirely plausible that the palatals arose from velars through the innovation of a process of palatalisation driven by an adjacent front high vocoid (eg, /i, e, j/), indeed palatalisation of this type is a very common phonological process. Szemerényi's recent (1996) IE Handbook concludes that "[m]ost scholars see themselves rather as forced to the conclusion that the palatals arose secondarily from fronted velars" (1996, 146); as Lehmann (1993) points out, this is already assumed by Meillet (1937).

GCS_{1b}
 p > f
 t > þ
 k > χ
 k^w > χ^w

GCS₂
 bh > ʙ
 dh > ɖ
 gh > ɣ
 gh^w > ɣ^w

GCS₃
 b > p
 d > t
 g > k
 g^w > k^w

The symbols used here are essentially those of Krahe (1969) (and of Prokosch 1933, Penzl 1975, König 1978, Sonderegger 1979, Schmidt 1984). We will not focus on the question of the independence or interconnectedness of these processes at this juncture, but we can note, with Lass (1997, 244), that the three main processes as set out here could logically have been independent, occurring in a range of chronological orders (although it appears that GCS_{1b} must have preceded GCS₃ because the output of GCS₃ seems precisely the same as the input of GCS_{1b} which would have lead to the segment classes merging if they occurred in the opposite order).⁷ While most of the reinterpretation of these processes will be saved to later chapters, some preliminary points can be made here. These are that the IE segment /b/ is very infrequently reconstructed; and the *TAs*, too, are never proposed to be frequent segments in IE. Both of these points are discussed in section 2.3.

⁷ This kind of argument has been used to propose ‘chain shift’ analyses for the whole GCS (see, for example, Martinet 1955) which seek to link the separable processes causally so that one process ‘set off’ another in order to maintain an underlying contrast, because its output segments started to become identical with already existing underlying segments (a ‘push chain’), or by ‘freeing up’ phonological space so that another process could be innovated to produce segments of a type which had been changed by a preceding process (a ‘drag chain’). The argumentation in Lass (1997), which merely notes that contrasts were not lost, does not commit us to a chain shift analysis, however. I discuss these points further in section 3.1.6.

Despite the impression sometimes given to the contrary, these changes were not strongly unconditioned (in the sense developed in section 1.2.3).⁸ Certain classic ‘exceptions’ to the changes have long been recognised, which show that the processes did not affect every occurrence of the IE underlying segments given in (2.1) and (2.2). I do not refer here to the factors which chronologically predated the GCS and hence prevented the IE *TAs*, *Ts*, *MA*s and *M*s from being potential inputs to the processes in (2.2), or factors which subsequently altered the outputs. Such factors have been recognised (eg, ‘Grassmann’s Law’ and ‘Verner’s Law’, see Collinge 1985). These are not discussed here, although some aspects of Verner’s Law are touched upon below. These factors would indeed affect the processes in (2.2) if they were taken to be simple correspondences between IE and Germanic, but they are not. They are interpreted as phonological processes and Verner’s and Grassmann’s Laws were not part of the actual input to the GCS processes themselves.

True ‘exceptions’ to GCS processes are generally only recognised for GCS_{1a} and GCS_{1b}. All the handbooks are explicit on this point, giving a precise set of phonological environments where the input to these processes were excluded from undergoing the spirantisation. The effect is generally formulated along these lines: “[the changes] generally avoided producing the situation where two fricatives would be next to each other. So IE *sp*, *st*, *sk* were never shifted and the IE clusters *pt*, *kt* only show a shift in the first stop, while the second remains unaffected” (Kienle 1960, 71, my translation).⁹ In this way, none of the IE *TAs* or *Ts* were changed to fricatives following /s/ or /p/ or /k/, to the extent that they occurred there in IE. Some of these exceptions will take on quite some importance later. Once GCS₃ had occurred, the exceptions to GCS₁ just mentioned merged with the new output segments which were created by the lexicalisation of GCS₃ and which are typically transcribed with the same symbols as IE *Ts*.

We might also note here that the strongly unconditioned formulation of GCS₂ as assumed in this section following most handbooks, is only the ‘majority position’ on the issue, as Vennemann (1984) describes it. A distinctly ‘minority’ position holds that there were

⁸ This is not always immediately obvious in the Handbooks and Grammars as it is not conventional to link the ‘structural change’ of a process to its ‘structural description’ in the philological literature. ‘Exceptions’ and concerns of the phonological environment of a process are typically given separately from the actual change, a practice which I adopt here.

⁹ The original reads: “Generell vermieden wird das Entstehen zweier unmittelbar aufeinanderfolgende Spiranten. Daher bleibn idg. *sp*, *st*, *sk* stets unverschoben und die idg. Lautgruppen *pt*, *kt* zeigen nur Verschiebung des ersten Verschlusslauts, während der zweite erhalten bleibt.”

exceptions, but we will not investigate this further here. We will return to the point later, however.

2.1.1.2 Evidence for the GCS

Some evidence for the GCS processes is given in (2.3), for all those cases where there is clear evidence to support it.¹⁰ This is comparative orthographic evidence, of the type which is conventionally cited in the Handbooks. It naturally only really becomes evidence for the processes once we interpret it in that way, as discussed at the start of this chapter. It can be interpreted thus: after the IE segments in the first column, a word is given in the second column from an IE language where the GCS processes were not innovated, and then reflex of the word in a Germanic language is given in the third column (which naturally shows evidence of the processes). The non-Germanic languages do not necessarily show the IE consonantism because they may have innovated different diachronic processes. The forms are orthographic as these are the key evidence that we have for the changes. To direct the reader's attention, the letter which represents the segment in question in the spelt form of the word is given in bold type, a convention which I also adopt elsewhere in this thesis.

¹⁰ The data is taken variously from Paul (1916), Kienle (1960), Krahe (1969), Lockwood (1969), Schmidt (1984) Prokosch (1939), Wright (1910). Kienle (1960) says that clear examples are lacking for IE /kh, khʷ/.

(2.3)

GCS_{1a}

ph	Old Indic	<i>phenas</i>	OHG	<i>feim</i>	'foam'
th	Greek	ασχηθηξ	Gothic	<i>skapis</i>	'undamaged, damage'

GCS_{1b}

p	Latin	<i>pes</i>	Gothic	<i>ƿotus</i>	'foot'
t	Latin	<i>frater</i>	Gothic	<i>broƿar</i>	'brother'
k	Latin	<i>canis</i>	Gothic	<i>hunds</i>	'dog'
k ^w	Latin	<i>sequor</i>	Gothic	<i>saihvān</i>	'see'

GCS₂

bh	Old Indic	<i>nábhās</i>	Old Saxon	<i>nebal</i>	'mist, cloud'
dh	Old Indic	<i>rudhirás</i>	Old Norse	<i>rauðr</i>	'red'
gh	Latin	<i>hostis</i>	Gothic	<i>gasts</i>	'enemy, foreigner'
gh ^w	Greek	ομφη	Old Norse	<i>syngva</i>	'voice, sing'

GCS₃

b	Old Bulgarian	<i>slabu</i>	Gothic	<i>slepan</i>	'limp, sleep'
d	Latin	<i>domus</i>	Old Saxon	<i>timberian</i>	'house, build'
g	Lithuanian	<i>áugu</i>	Old English	<i>eaƿian</i>	'grow, multiply'
g ^w	Old Indic	<i>gam-</i>	Gothic	<i>qiman</i>	'come'

The data in (2.4) shows evidence for the effect of a preceding /s/ or /p/ or /k/ on the *Ts* of Indo-European, which remain *Ts* in Germanic (as, among others, Krahe 1969 explains, there are no clear examples for /k^w/).

(2.4) Exceptions to GCS_{1b}

p	Lithuanian	<i>spiauju</i>	Old Frisian	<i>spiwa</i>	'spit'
t	Latin	<i>stella</i>	Old English	<i>steorra</i>	'star'
	Latin	<i>captus</i>	OHG	<i>haft</i>	'captured, capture'
	Latin	<i>octo</i>	Old Frisian	<i>achta</i>	'eight'
k	Lithuanian	<i>skabù</i>	Gothic	<i>skaban</i>	'cut, clip'

2.1.1.3 Where and when did the GCS occur?

Both Schrodtt (1979) and Collinge (1985) comment on the wide range of datings that have been proposed for the GCS. One obvious problem in giving even a vague date for the changes

is that the four parts discussed here (1a, 1b, 2 and 3) need not be seen as parts of a unified process. While some have proposed this, most authors countenance their being separable, even if it is proposed that the changes are indirectly related in that the occurrence of one laid the ground work for another. It is also possible to propose that the occurrence of one of the changes caused one or more of the others to be introduced, particularly in 'struturalist' approaches. Schrodts (1979) shows that the dates that have been proposed for the GCS by scholars vary by up to two millennia, from around 1500 BCE into the first millennium CE. Different types of evidence (the form of loanwords in borrower and lender languages, the form of early inscriptions, archaeological factors) lead to different conclusions, but after surveying the available proposals, Schrodts (1979) settles on around the fifth century BCE (allowing a couple of centuries leeway). This would probably be in accordance with most scholars' opinions on the issue and it genuinely does not seem to be possible to be more precise; we can, however, presume that the processes occurred at some point when, and in some place where the Germanic peoples were still closely in contact with each other.

Schrodts (1979) also proposes that GCS_{1b} was the first process chronologically. It is likely that we would want to link GCS_{1b} with GCS_{1a} because the output is the same, and Lass (1997) shows that it is reasonable to assume that GCS₃ followed this, given the formulations of the processes that we have adopted here. As Lass (1997) also points out, we can make no certain claim as to the relative dating of GCS₂. The precise data of the changes is not important for this thesis, so we will not pursue the matter further; what does matter is that something along the lines of that given in (2) did happen. While we can indeed be sure that some phonological processes were introduced between IE and Germanic, we will see later (in sections 2.3 and 4.1.1) that distinct doubts have been raised about the formulation of the GCS as given in (2.2). For the moment, however, we will proceed with the standard, Handbook-style formulation.

2.1.1.4 Is the GCS an example of lenition?

Given that the focus in this thesis is on the notion of lenition, it will be well at this juncture to consider whether the processes discussed here can be seen to count as lenitions, according to the preliminary definition adopted in section 1.4. Once we have returned to the changes in chapter 4 to reanalyse them, it will be clearer, but even here, we can see that some of the key processes of the GCS fit squarely onto 'lenition trajectories' of the type discussed in section

1.4. Both GCS_{1a} and GCS_{1b} have voiceless stops as their input and voiceless fricatives as their output, as shown here in (2.5) which repeats parts of (2.2) for convenience:

(2.5)

GCS_{1a}
 $ph > f$
 $th > \theta$
 $kh > \chi$
 $kh^w > \chi^w$

GCS_{1b}
 $p > f$
 $t > \theta$
 $k > \chi$
 $k^w > \chi^w$

The major place of articulation is not affected in any of the segments and it seems that the data here can be described as a straightforward case of ‘opening’ lenition of the type stop \Rightarrow fricative, and it has been described overtly in this way (eg, in Lass 1984, Labov 2001).

The formulation of GCS_2 given in (2.2), repeated here as (2.6), with the symbols θ , δ , g , g^w representing fricatives, also seems to be a relatively straightforward lenition: the inputs are stops (of some sort), the outputs are fricatives, and the change does not affect the segments’ place specifications. This, too, appears to be a clear case of ‘opening’ lenition of the type stop \Rightarrow fricative:

(2.6) GCS_2

$bh > \theta$
 $dh > \delta$
 $gh > g$
 $gh^w > g^w$

Against this, GCS_3 , repeated here as (2.7), seems quite clearly not to be a case of lenition, given the definition in terms of trajectories and process-type in section 1.4.

(2.7) GCS_3

$b > p$
 $d > t$
 $g > k$
 $g^w > k^w$

Given that GCS₃ is typically grouped as part of the GCS, it is interesting to consider what, if any, relationship it has with the other ‘parts’ of the GCS, a temptation which few theorists have been able to resist. I turn to this question in section 4.1.1.

2.1.1.5 An appendix: ‘GCS’ debuccalisation

On considering the type of IE-Germanic correspondences which served as the basis for the formulation of the GCS, as shown (2.3), several points stand out. It will be helpful for the development of later arguments to discuss one of these here, before we leave the realm of the GCS. This point is closely connected with the data described in this section as it derives from the subsequent fate of some of the segments which resulted from the change. It also represents a further case of lenition. I focus here one segment in particular: that which has been transcribed as Proto-Germanic /χ/ (on the picture given here, this resulted from IE /k/ and /kh/). The same basic point applies to the labiovelar /χʷ/ (from /kʷ/ and /khʷ/), but we shall focus on the simplex segment.

The evidence given in (2.3) for this case makes the point: Latin *canis* ‘dog’ corresponds to Gothic *hunds*. While Latin <c> quite clearly represents /k/, unchanged from IE, the Gothic <h> most likely represents a glottal /h/, rather than the oral fricative /χ/.¹¹ This is clearly the case in Modern Germanic languages (eg, English /haund/ *hound*, Dutch /hond/ *hond*, German /hund/ *Hund*). This correspondence exists in word-initial environments and is traced back to common Germanic (see, for example, Luick 1914-1940, 834). The same situation is thought to have held for the environment between two voiced segments.

It seems entirely reasonable to posit a fricative stage before the reflex of IE /k/ (and /kh/) became /h/ because the reflex of IE /k/ in many phonological environments, other than word-initial, is clearly a dorsal fricative in most non-English present-day Germanic languages (for example, in Dutch *nacht* and *licht*, German *Nacht* and *Licht* ‘night’ and ‘light’ - compare Latin stems *noct-* and *luc-*) and there is good orthographic, orthoepic, dialectal and diachronic evidence that it existed in earlier stages of English, too (see, for example Lass 1997, 220). The

¹¹ It may seem like the use of philological symbols in this chapter may be starting to become problematic here as the question arises whether the symbol ‘χ’ represents a velar (like the input stop) or a uvular (in accordance with the value of the symbol in the IPA). Picard (1999) comments overtly on this point and proposes that the use of the symbol χ in philological texts could be responsible for confusion about the precise place of articulation of the fricative. However, the situation is more complex than Picard (1999) makes out. The segments transcribed as /x/ in modern Germanic languages, at least, are typically uvular, not velar (see Kohler 1977, 1990 and Honeybone 2000a, b). Given this, the philological symbol is perhaps usefully vague and I shall not deviate from using it here. We return to this point in section 2.3.

orthoepic evidence further suggests that this fricative went through an /h/ stage before being lost in English and it seems reasonable to extend this back to the earlier common Germanic process to conclude that this is a case of lenition by debuccalisation. The full story for IE /k/ in terms of the GCS and beyond is thus a trajectory of the type given in (2.8):

(2.8) stop \Rightarrow fricative \Rightarrow glottal

$k > \chi > h$

This change $\chi > h$ is generally reckoned to have occurred in Germanic “very shortly” (Krahe 1969, 98) after GCS₁. This debuccalisation (along with the preceding spirantisation) fits neatly on the lenition trajectories discussed in section 1.4 and qualifies as an example of lenition in its own right.¹² This is not intimately connected with the other changes discussed here, in that sense it is not a part of the GCS. It is merely the case that the input to the change was created by a GCS process. But it has already been argued that the individual changes discussed here as ‘parts of’ the GCS are not necessarily intimately connected with each other. They will prove important in the discussion in chapter 3, however, as will debuccalisation of the type discussed here.

2.1.2 The High German Consonant Shift

The second set of data, to be discussed in this section, is also well-known in philology, especially in Germanic philology, and is also discussed, although much less so, in theoretical linguistic literature. The segmental changes dealt with here under the label ‘High German Consonant Shift’ characterise, as the name suggests, the ‘High’ dialects of modern German (that is, approximately, those with a geographical base in the middle and south of the German speech continuum, see, for example Paul 1916, Russ 1990). As we will see below, present-day surviving traditional dialects show a difference in the extent to which they exhibit the segmental changes which were the result of the introduction of the HGCS phonological processes. Indeed, the HGCS is the classic reference point of the study of German dialectology, as it is standardly taken to be the basis for the distinction between the two largest indigenous Germanic dialect-groupings (or, maybe, languages) spoken in modern Germany:

¹² As is well-known, in many varieties of English /h/ has subsequently been lost, so a further stage can be added to the trajectory: $k > \chi > h > \emptyset$.

Low German and High German.¹³ Those varieties which feature underlying segments which are lexicalisations of the HGCS processes are High German, those which do not are Low German. Needless to say, I shall have nothing to say about Low German varieties in this section.

The phenomenon in question here is also known as the ‘Second Sound Shift’ (*die zweite Lautverschiebung*), especially in literature in German on the topic, although the more descriptive ‘HGCS’ is also usually given (as in Paul 1916, Vennemann 1984). This derives from the common perspective, which is accepted unquestioningly in my exposition here, that it followed chronologically after the GCS (which is the ‘First Sound Shift’ in such terminology, as mentioned above). Such a title might be seen to unduly emphasise its importance, however. It is a substantial change, but not on a scale never attested in other languages (an at least partial analogue is described in section 3.2.1.8).

2.1.2.1 The HGCS: take 1

This section will deal with a range of issues that are connected with the HGCS and will again feature presentation which is in a largely philological style. One key question in connection with the set of phenomena which can be brought together under the name ‘High German Consonant Shift’ is whether they are to be conceived of as one process or as many. We will return to this question in section 4.2 and so will not devote too much space to it here, but the demands of presentation will require us to make a distinction between two putative separate processes. This is conceptually similar to the division of the GCS into separate processes, and a distinction of this sort is commonly made in the literature (Braune 1891, Paul 1916, Keller 1978, Vennemann 1984, Davis, Iverson & Salmons 1999) with some authors claiming that one of the processes is the ‘real’ HGCS and the other a later, much less important development.

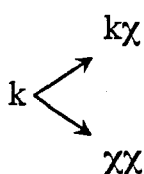
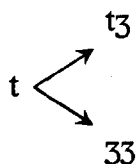
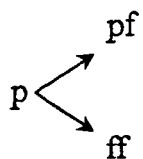
The distinction I make here is in terms of the input to the processes. The best known and

¹³ A description of the situation like this cannot help but raise a large number of questions which are well worth pursuing in their own right, but which would unfortunately take us too far afield here. Is Low German a separate language from High German, or is it a ‘dialect of German’? Is it still spoken widely enough to be described as ‘spoken in modern Germany’? Is it reasonable, on historical linguistic grounds, to speak of High German as one variety anyway? Isn’t Frisian an ‘indigenous Germanic language spoken in modern Germany’? These questions have long been discussed in the literature. The answers are not so clear, however, and I adopt what I take to be a relatively standard position here.

most discussed process affected the Proto-Germanic *Ts* /p, t, k/ (/k^w/ had been lost as a unary contrasting segment on the way from Proto-Germanic to pre-HGCS Old High German). These were the output of GCS₃ and the process which affected them is the part which Vennemann (1984) singles out for description as the ‘HGCS’ by itself; it is also the only process that Davis and Iverson (1995) discuss under the heading ‘HGCS’, all of which might well lead us to conclude that it is separable from the second process which is sometimes claimed to be part of the HGCS. I label it HGCS₁. The second affected the Proto-Germanic *Ms* /b, d, g/ (/g^w/ had also been lost). These segments were etymological reflexes of the output of GCS₂, because it is typically claimed that the fricative outputs of GCS₂ had become stops in pre-shift OHG (Prokosch 1933 even gives this pre-process a name: ‘the Intermediate Shift’). The second process is sometimes referred to in German as the *Medienverschiebung*, where *Medien-* is a reference to the type of segment which we have referred to in places as ‘Mediae’ (*Ms*). Davies, Iverson & Salmons (1999) call it ‘the Shift of the Voiced Stops’. I label it HGCS₂.

In the traditional philological rendering of HGCS₁, the vast majority of Germanic *Ts* became either affricates or geminate fricatives, whilst maintaining their basic place of articulation. I turn to what conditioned this split imminently, but first, (2.9) shows a simple way of representing the diachronic correspondences that were created and which can in part still be observed by comparing modern German with other Germanic languages (the arrows have shafts here only for the ease of presentation). The segments on the left of the arrows is the reconstruction of the Proto-Germanic stop inventory which is commonly assumed in the ‘Germanic’ Handbooks (eg, Braune 1891, Paul 1916, Moulton 1972, Ramat 1981) minus the labiovelars:

(2.9) HGCS₁



The symbols used here for the outputs are meant as a distillation of those used in philological work. They are almost exactly those of the authoritative Moulton (1954). I also follow Moulton (1954) and others in presenting the most general version of the processes here; this occurred only in southerly varieties of High German (Alemannic and Bavarian). The symbol ‘ʒ’ is a compromise which is often used in Handbook-style treatments at least partly in the way that I use it here (for example, in Paul 1916, Braune 1961, Sonderegger 1979).¹⁴ It represents a voiceless fricative and is certainly not meant to be interpreted with the value of IPA [ʒ]. This is not entirely consistent with the practice of OHG scribes; in the manuscripts, both segments were often represented with the same spelling, sometimes simply <z> (although some OHG scribes did make a consistent difference, for example, as Penzl 1970 explains, in the *Isidor* manuscript the affricate is systematically written as *z* or *tz* and the fricative always as *zs*). An important point is made by the use of ‘ʒ’ in the Handbooks: the fricative output of the HGCS₁ process did not merge immediately with the already existing Germanic fricative, which is typically written as <s> and which was inherited unchanged from Indo-European. We will return to this transcription and the question of exactly what the symbol represents in section 3.1. It is clear, however, that ‘tʒ’ represents a voiceless coronal affricate and ‘ʒʒ’ a voiceless coronal geminate fricative.

¹⁴ Paul, for example, uses ʒʒ for the fricative against *z* for the affricate, while Braune (1961) uses ‘ʒ’. Earlier editions, such as (1891) simply use ‘z’ for both affricate and fricative (although the latter is doubled: ‘zz’).

The traditional account for the split in the realisation of the Proto-Germanic segments (and subsequent underlying split) proceeds as follows: stops became affricates if they were in word-initial position or occurred directly following a consonant (including the first half of a geminate; Proto-Germanic had developed quite a large number of geminates in the ‘West-Germanic Geminatation’, widely described in the Handbooks, for example where segments occurred at the end of a syllable word-internally before /j/). In contrast to this, fricatives developed if the original stop occurred postvocally.

One way of understanding this difference is to unify the processes, so that the fricatives first went through a stage of affrication before further change to fricatives (and this has been proposed in the literature). Due to the fact that the HGCS occurred before the earliest OHG texts, there is no direct evidence for this stageist hypothesis, but it is possible to conceive of the phenomena as one general process which was partially inhibited in the environments which now show affricates. In any case, the processes were entirely inhibited in certain environments. Exceptions to HGCS₁ are recognised for Germanic /p, t, k/ such that the stops remain as stops when they formed part of consonant clusters: /st/, /sp/, /sk/, /ft/, /ht/, /tr/. We will return to these below, but we can note here that it is typically argued that the exceptions to HGCS₁ merged with the outputs of HGCS₂ lexicalisation.

As already mentioned, HGCS₂ is less widely discussed than its co-labeller. It is typically thought to be both “later and more geographically restricted” (Davis, Iverson and Salmons 1999, 192). Evidence can be adduced from the comparison of the orthography of NHG with that of OHG, that the alveolar *M*, /d/, was affected in all positions in the varieties which went on to form the basis of the standard language, and was accompanied by analogous developments to /b/ and /g/ (forming a natural class) in southern dialects. This notion of the relation between ‘standard’ and ‘dialect’ is problematic, however, as was discussed in section 1.2.3. In terms of symbols, the following is generally given for an extreme variety:

(2.10) HGCS₂

b > p

d > t

g > k

It seems from the orthographic record that, in at least some varieties, this occurred across the board, in all phonological environments, although the orthographic evidence is not entirely

straightforward to interpret because it is often the case that there is alternation in the spelling of the segments, so that the labial, for example, is sometimes spelt <p> and sometimes spelt .

2.1.2.2 Evidence for the HGCS

The columns of (2.11) contain some orthographic evidence for the HGCS processes.¹⁵ In a similar style to the presentation in section 2.1.1, a word in a Germanic language where the HGCS processes were not innovated is given in the second column and the reflex of the word in a High German variety in the third column. The first column contains the Germanic form of the segment concerned. Evidence is provided for all the processes given in (2.9) and (2.10) and at times this necessitates giving data from the most southerly, ‘Highest’ or Upper German varieties. For the HGCS₁ processes, the first row gives an example of an OHG word with an affricate, the second a word with a fricative:

(2.11) HGCS₁

p	Old English	<i>pund</i>	OHG	<i>pfund</i>	‘pound’
	Old Saxon	<i>opan</i>	OHG	<i>offan</i>	‘open’
t	Old Saxon	<i>settian</i>	OHG	<i>setzan</i>	‘sit’
	Old English	<i>etan</i>	OHG	<i>eʒzan</i>	‘eat’
k	Old Saxon	<i>werk</i>	OHG	<i>uuerccch</i>	‘work’
	Gothic	<i>wakan</i>	OHG	<i>uuahhen</i>	‘wake’

HGCS₂

b	Old Saxon	<i>dohter</i>	Upper OHG	<i>tohter</i>	‘daughter’
d	Old English	<i>beran</i>	Upper OHG	<i>peran</i>	‘bear’
g	Gothic	<i>giban</i>	Upper OHG	<i>keban</i>	‘give’

The data in (2.12) shows evidence for the effect of adjacent consonants on the Proto-Germanic *T*s, which would otherwise be inputs to HGCS₁:

¹⁵ The data is taken variously from Braune (1991), Paul (1916), Schmidt (1984) .

(2.12) Exceptions to HGCS₁

p	Old English	<i>spinnan</i>	OHG	<i>spinnan</i>	'spin'
t	Gothic	<i>stains</i>	OHG	<i>stein</i>	'stone'
	Gothic	<i>nahts</i>	OHG	<i>naht</i>	'night'
	Gothic	<i>luftus</i>	OHG	<i>luft</i>	'air'
	Old English	<i>treowe</i>	OHG	<i>ttriuuua</i>	'true, truth'
k	Gothic	<i>fisks</i>	OHG	<i>fisk</i>	'fish'

2.1.2.3 Where and when did the HGCS occur?

It has already been mentioned in this section that the processes described here were innovated to their fullest extent in southerly, Upper German dialects; as we also saw previously, no part of the HGCS occurred in northern, Low German dialects. It is most commonly argued on this basis (for example by Paul 1944, Sonderegger 1979, Schmidt 1984 and many others) that the processes were first innovated in the South and subsequently spread northwards. This is thought to fit in with the range of isoglosses which exist in the present-day traditional dialects of High German such that, for example, the further north a dialect is, the less phonological environments there are where the processes have occurred. It is also the case that a difference exists as to which segments were affected by the processes. For example, in HGCS₁ in the northernmost dialects (ie, in certain varieties of Central German such as Central Franconian), the stop /t/ was affected in more environments than were /p, k/, and in South Rhine Franconian, where /p/ was more affected by the process than in Central Franconian, it was not affected in word-initial position (Sonderegger 1979).

The basic idea behind the standard position is that the processes originated where they affected most segments in most environments and gradually spread northwards 'losing force' as they went. Some have argued the opposite, namely that the process spread from north to south, becoming more general as it went (for example King 1969 who bases his claim on a Standard Generative Phonology model of rule complicatedness and the assumption that rules are simplified as they spread). There is, however, also an argument from the investigation of surviving manuscripts which places the earliest recorded evidence for the processes in the South. It is not unduly surprising that a process should be slightly different in dialectically different areas. Given the points made in chapter 1, however, we would expect the various formulations to be 'possible' changes to the extent that they are not imposed from outside.

For later purposes, it will be worth briefly investigating the precise dialectal developments in the better-known HGCS₁. Keller (1978) provides a perceptive summary. He recognises three ‘contexts’: stops in context 1 are in the environments which were identified in (2.12) above as exceptional; this context inhibits the processes in all dialects. Context 3 is defined as “medially and finally after vowels” (Keller 1978, 169); the processes affect “all Upper German and central German dialects in context 3” (1978, 171). Context 2 is “initially, after liquids, and nasals, ... in gemination” (1978, 169); it is here that the greatest differences can be found. Keller presents the inhibitory effects of specific environments in the various dialects in tabular form, reproduced below as (2.13).¹⁶

(2.13)

	Highest Alem.	H. Alem.	L. Alem	Mos. Fr. Rh. Fr.	Rip.	L. Franc	L. Sax
<i>Zinn</i> ts-							t- tin
<i>setzen</i> -ts-							-tt- setten
<i>Wasser</i> -ss-							-t- water
<i>Schlafen</i> -f, -ff-							-p- slapen
<i>machen</i> -ch-							-k- maken
<i>werfen</i> rpf > rf							-rp werpen
<i>helfen</i> -lpf > lf							-lp helpan
<i>Dampf</i> -mpf							-mp Damp
<i>Apfel</i> -pf-							-pp- Appel
<i>Pfund</i> pf-							p- Pund
<i>starch</i> -rkx > rch							-rk stark
<i>mälche</i> -lkx > lch							-lk melken
<i>Chind</i> kx > ch-							k(h)- Kind
<i>teiche</i> -ŋkx > (ŋ)ch-							-nk denken

¹⁶ To expand Keller’s abbreviations: Highest Alem. = Highest Alemannic; H. Alem. = High Alemannic; L. Alem = Low Alemannic; Mos. Fr. = Mosel Franconian; Rh. Fr. = Rhine Franconian; Rip. = Riparian; L. Franc. = Low Franconian; L. Sax = Low Saxon. The italicised example words are present-day standard German, the others are dialectal.

From this we can see that Keller's context 3 was entirely uninhibitory in Highest Alemannic, but that, for example, in High Alemannic, /k/ was not affected in the sequence /ŋk/. The evidence presented here is often used to argue that the HGCS₁ originated in the most southerly, High Alemannic dialects.

Similar, although not quite identical effects are claimed for HGCS₂. For example, Sonderegger (1979) places the centre, where the greatest effects can be seen, somewhat further north than for HGCS₁ and claims that this is where the process has its origins.

As with the GCS, various suggestions have been made as to the dating of the HGCS and the various dialectal developments just described, but there is not as much variation in the proposed dates. This is because, at least in the formulation presented here, it must follow the GCS. It is also reasonable to assume that it followed a substantial wave of geographical spread and separation of the Germanic peoples. The major change, HGCS₁, must also precede the composition of the earliest OHG texts, which can be approximately dated to the middle of the 8th century by various means, because attempts to represent it are apparent in the orthography of these manuscripts. Some other classic pieces of evidence lead Penzl (1970), and Wolf (1981), among many others, to date it to around the 6th or 7th century CE. The HGCS₂ is generally placed a century or two later.

Some recent proposals have called a substantial proportion of the type of presentation given here into doubt. Key among these are Vennemann's (1984, 1992) writings on the topic. We investigate some of the points made there in chapter 4, but we may note here that he proposes that the HGCS (at least, HGCS₁) occurred much earlier than is generally assumed. Vennemann (1984) presents a reinterpretation of some of the occurrences of the HGCS and Vennemann (1994) argues for a date for these of around the first century BCE. The precise dating of the processes is not of substantial importance for our purposes, however; what is important is that something along the lines of that described in this section did in fact occur. While Vennemann's reformulation is notable, even if it is correct, we may be reassured, with von Stechow (1986), that it is not so substantial as to entirely move the goalposts. Vennemann (1994) proposes that much of the dialectal situation, described briefly above, is illusionary and really represents patterns of repression on the lexicalised results of the processes. This would mean that the original, natural change throughout the High German speech community was the most extreme version, currently found in High Alemannic, with all three affricates, for example. This is an interesting, if controversial, proposition, but even if it is true, it will make

little difference to the coming discussion; one point to note is that this does still not make the process strongly unconditioned because even in the most fully affected varieties, there were still some ‘exceptions’. For the moment, we will proceed with the standard, Handbook-style formulation.

2.1.2.4 Is the HGCS an example of lenition?

After the extensive discussion of the HGCS processes in this section, especially of HGCS₁, the question arises as to whether these processes are characterisable as examples of lenition. At first glance, this is not obviously the case. The process described in HGCS₂, and repeated here as (2.14), seems clearly not to be a lenition. Like GCS₃, it represents a backwards step on the trajectory voiceless \Rightarrow voiced from section 1.4.

(2.14) HGCS₂

b > p

d > t

g > k

The processes of HGCS₁ seem much more lenition-like but there is some problem with fitting them squarely on a simple lenition trajectory. One problem is that the result of the processes was a segmental split; another is that, while aspects of the processes look like a case of stop \Rightarrow fricative, the fricative is a geminate.¹⁷ There are two ways of conceiving of HGCS₁ which make it more tractable. These are given in (2.15) and (2.16).

(2.15) HGCS₁ as two one-step processes

p > pf

t > tʒ

k > kχ

p > ff

t > ʒʒ

k > χχ

¹⁷ As explained in section 1.4, I assume with among others, Anderson & Ewen (1987) and Ségéral & Scheer (1999) that stop \Rightarrow affricate is a clear case of ‘opening’ lenition. This is further discussed below.

(2.16) HGCS₁ as one two-step process

p > pf > ff

t > t₃ > ʒ

k > k_χ > χχ

The conception in (2.16) especially makes it seem reasonable to view the HGCS as consisting of subsequent (or even simultaneous) stage 1 and stage 2 lenition. However geminates are understood, the output geminates *are* fricatives so it does not seem to be an unreasonable preliminary assumption that they might be further along a lenition scale than the original stops. The place of articulation is not affected in any of these changes and so we might describe the data here as a case of ‘opening’ lenition on a trajectory stop ⇒ affricate ⇒ fricative. However, where such concepts are discussed, it is generally argued (for example in Foley 1977 and Bauer 1988) that geminates are phonologically stronger than simplex segments. This poses a real problem for our understanding of the HGCS, to which we return in section 4.2.

2.1.3 English Initial Fricative Voicing

While it cannot compete for fame with the GCS and the HGCS, the process which I describe in this section is quite well known among English philologists and historical linguists. The fricative segments of English have quite a complex history and their development has generated some substantial discussion. The phenomena that I deal with here do not have a monolithic label, unlike those discussed above, which have several. It has been referred to as ‘Old English Fricative Voicing’ (for example, by Lass 1991-93) and also as ‘the Voicing of Initial Fricatives in Middle English’ (by Fisiak 1984, for example), and often without any special designation. As some name is needed, I use the phrase ‘English Initial Fricative Voicing’ which reflects the fact that the effects of the processes in question come to light in the historical orthographic record of English; they can also, in fact, be heard in present-day non-standard dialects.

Due to the ambiguity of the word ‘voicing’ in English, it is not clear from Lass’s and Fisiak’s titles whether they intend the term to be taken dynamically or statively (does ‘voicing’ refer to a process or a state?) and this is the probable explanation for the apparent difference in dating in their titles. As we will see, there are real issues surrounding the question of precisely when the processes discussed here occurred and it has been situated as part of Old English, Middle English and other stages of the language. Irrespective of the label, the developments

merit a mention in Handbook-style treatments of English historical phonology, such as Ellis (1869), Sweet (1888), Wright (1928), Luick (1914-1940), Mossé (1952), Brunner (1965), Prins (1972) and Lass (1992).

There is no doubt that the data show that some type of process has been innovated, but it is also clearly the case that the process did not affect all varieties of English. Coincidentally, like the HGCS above, the effects of EIFV are only in evidence in southern traditional dialects and, because these dialects did not form the basis of the standard language, there is little evidence of the process in standard English orthography or in Northern English and Scottish accents, for example, or in the RP ‘reference’ accent; the only words in these varieties which feature segments affected by the process are borrowings from southerly varieties.

2.1.3.1 The EIFV: take 1

It is generally recognised that there was only one series of underlying fricatives in all varieties of Old English. It is normal in Handbook presentation to distinguish between two separate phenomena in dealing with the laryngeal states of these fricatives. The normal position is to describe one of these as the original Old English situation, which was sustained in Northern and Midland dialects, and the other as a Southern innovation. It is the second of these which we are dealing with here as EIFV.

In probably all varieties there were two types of surface fricative segments, standardly described as voiced and voiceless. Because it is relatively clear that these were in complementary distribution (although see Bammesberger 1988 for some dissenting comments), it is unsurprising that this has generally been described as a classical case of allophony (even before the phoneme was an overt part of phonological theory the same basic story was told). It is difficult to describe this insightfully without using some type of phonological terminology, so I diverge slightly from the manner of presentation adopted in previous sections in order to make overt what is implied in older Handbooks. In fact, it is highly likely that at least some of the segments which were characterised so simply above in the discussion of the GCS and HGCS were in fact subject to some type of phonological processes which gave them more than one surface realisation. This is not easily recoverable from the evidence that we have, however. The realisation rule given here for EIFV is well established and well-known, though, so I include it below for clarity.

If we focus on the phonology of original, non-Southern Old English, it is typical, where some type of phonemic theory is employed (as in Hogg 1992c and Lass 1994), to describe the underlying segments as being voiceless and to assume that there was a process whereby these underlying segments were realised as voiced surface segments medially. Lass (1994, 72) formulates this as a phonological rule, reproduced here in (2.17).¹⁸ Lass uses IPA [θ] where previously in this chapter the philological ‘þ’ has been used, and V stands for any vowel and (L) for an optional liquid (the first vowel also carried stress):

(2.17) Non-EIFV Old English:

$$\begin{bmatrix} f \\ \theta \\ s \end{bmatrix} \rightarrow \begin{bmatrix} v \\ \delta \\ z \end{bmatrix} \quad / \quad V \text{ (L) } __ V$$

This is an example of a common type of voicing process. While not all examples of the voiced surface segments were originally derived by the innovation of this process, as Hogg (1992c) and Lass (1994) explain, many of the them were, and it is clear that at some point in the history of English the process was innovated. Luick (1914-1940) in fact describes this as a common Germanic process, which can also be seen to have occurred in Old Frisian, Old Saxon and Old High German, so the innovation of the process can probably be dated before the Anglo-Saxon invasion of Britain. While this process is easily describable as a lenition, it is not the focus of this section and we will not investigate it further (for a detailed discussion see Lass & Anderson 1975). To return to the true focus of this section, the process in (2.17) can still be seen as part of the phonology of pre-EIFV Old English, which was the input to the EIFV. This input thus consisted of underlyingly voiceless fricatives which were contextually voiced.

The EIFV can be seen as changing this situation substantially. The standard position is that the process affected all the fricatives in the phonological system which could occur in the relevant environment, which as the name given to the process suggests, was domain-initial. The only exception is generally claimed to be the velar fricative, which, as we saw in section 2.1.1.1, had been lost in Germanic in this environment and had a very restricted distribution. If we follow Luick’s (1914-1940) formulation (and add in the segment /ʃ/ which was derived

¹⁸ Elsewhere, Lass (1992) describes it thus: “/f θ s/ were in most dialects voiceless except medially in the foot” (1992, 41) and see Lass (1991-1993) for detailed exemplification of the environments.

from Germanic /sk/ clusters and which Fisiak 1984 and Lass 1991-93 show also underwent the process), the EIFV can be represented as in (2.18):

(2.18) EIFV¹⁹

f > v

þ > ð

s > z

ʃ > ʒ

This process is thus claimed to have involved the voicing of segments which were underlyingly voiceless. It is typically described as having occurred domain initially, although it is not entirely clear that any voiceless fricatives remained in the varieties where this process was innovated. If they did, then fricatives would only be voiceless when adjacent to a voiceless segment and domain finally.

2.1.3.2 Evidence for the EIFV

It is not as straightforward to cite evidence for this process as it was for the GCS and the HGCS. This is partly because the EIFV process did not occur in varieties which have gone on to form the basis for standard languages, unlike the GCS and the HGCS. This means that the orthographic record is not as clearly helpful here as it is for the previous two processes. There is orthographic indication of certain aspects of the EIFV which dates from the Middle English period, when there was substantial diversity in English spelling, but even here the evidence is not copious, thanks to the spelling conventions that were available to writers in English at the time. As Lass (1991-1993) explains, the writers of OE had at their disposal either or <v/u> to represent a version of the labial fricative. If these letters were used in spelling EIFV words, this would be a clear change from the previous convention of using <ƿ>. A point that would encourage such spellings, if the speaker's variety had undergone the EIFV was that ME scribes had Latin as a model for a language with contrastive voiced fricatives which were written by <f> and <v/u>. Written forms are found featuring these unambiguously 'voiced' letters, rather

¹⁹ The formulations here revert to philological 'þ' where Lass used IPA /θ/. The symbols are a little problematic here if we do not use the IPA or another modern tradition, partly because the letters <þ> and <ð> were both used in Old English orthography, but were used relatively interchangeably (see Lass 1991-93, 6 for some generalisations about their usage). Naturally, the symbols here are intended to represent phonological segments, not letters.

than <ƒ>, in texts from the southerly dialects of Middle English. There was not the same possibility to indicate voicing for the other fricatives, however. The letter <z> was rarely used, and when it was used, it normally stood for then an affricate (except in the *Ayenbite of Inwyt* manuscript, as shown below). There was also no possibility of indicating voicing in the new palato-alveolar /ʃ/, nor in the inherited dental fricatives, where <þ> and <ð> were both used interchangeably by most scribes.

Some of the orthographic evidence that does exist is given in (2.19), taken from Luick (1914-1940). Middle English spellings are given which compare the use of <u> with modern standard English spellings with <f>, and spellings with <z> (from *Ayenbite of Inwyt*, where a distinction was made in terms of voicing, unusually, between <s> and <z>) are compared with modern <s> spellings. The first column gives the input, which were typically the Proto-Germanic output of GCS₁:

(2.19) EIFV

f	Middle English	<i>uader</i>	Middle English	<i>uram</i>
	Modern English	<i>father</i>	Modern English	<i>from</i>
s	Middle English	<i>zenne</i>	Middle English	<i>zuord</i>
	Modern English	<i>sin</i>	Modern English	<i>sword</i>

Some important evidence for the EIFV comes from the phonology of present-day non-standard traditional dialects. These show evidence that EIFV was active in these variety and it is reasonable to project this back to OE, because other evidence seems to suggest that the process was innovated then. Some of the evidence is presented in (2.20). This is fieldworker transcriptions of twentieth century non-standard dialects which were recorded for the Survey of English Dialects (see for example Orton 1962). The examples are lifted from Lass (1991-1993) who describes the relevant dialect area as “bounded on the north by a line running roughly from the southeast corner of Herefordshire down through Surrey and Kent (with a tiny island in east Sussex), and bounded on the south by a line running roughly through mid Cornwall” (Lass 1991-1993, 13). The forms in (2.20) show transcriptions of initial voiced fricatives where the standard has voiceless segments. The Germanic input segments are given in the first column:

(2.20) EIFV

f	[v] in <i>farmer</i>
þ	[ð] in <i>thumb</i>
s	[z] in <i>six</i>
sc	[ʒ] in <i>shilling</i>

2.1.3.3 Where and when did the EIFV occur?

Two questions arise when we consider the geography and dating of the EIFV. It is possibly easier to give a convincing answer to the first than the second. We can ask (i) which of the geographic varieties of English were affected and (ii) where the process was first innovated; and for each of these, we can ask when?

While we cannot be entirely sure of the precise details, Fisiak (1984) shows by means of an impressive array of dialectal, onomastic and manuscript evidence that the process affected the dialects of the whole of the South of England and parts of the West Midlands, but not East Anglia, or the East Midlands or any further North. It is clear from the 20th century dialectal situation just described that this area of influence has decreased in size in that there is currently no evidence of the lexicalised remains of the process (eg, initial voiced fricatives in words like *six* and *thumb*) in the current dialect of much of the South and West Midlands of England. A reasonable assumption is that pressures of prestige and standardisation have caused speakers in some of the original area of the EIFV to adopt more standard-like forms of English, which were not affected by the process. It is possible that the process affected the segments to differing degrees in different dialects, rather like the HGCS. In fact, this is quite often proposed given the difference in the extent to which the process can be represented in spelling. We cannot now be certain about this. The situation in modern dialects according to the SED is that the distribution of voicing is extremely similar for all segments, and Fisiak's (1984) evidence is consistent with a uniform spread, even taking in the late developing /ʃ/.

The fact that the process affected /ʃ/, the development of which segment both Luick (1914-1940) and Fisiak (1984) date to around the tenth century CE, might lead us to propose a similar or somewhat later date for the EIFV, as many have done (eg, Luick 1914-1940, Wright 1928, Brunner 1965). There are, however, in fact two distinct traditions as regards the answer to question (ii), and these have implications for the point under discussion here. The standard opinion, which I have allowed to guide the discussion in this section, holds that EIFV is a truly endogenous innovation

which began in Southern England around the aforementioned date. The other tradition, advocated by Bennet (1955), Nielsen (1981) and Lass (1991-1993), connects the process with at least partially analogous events in other Germanic languages. On such an account the EIFV either becomes a shared West Germanic innovation (Bennet 1955, Lass 1991-1993), which pushes the date much further back, or a process which is borrowed from closely proximate Germanic languages, through contact (a proposal which Nielsen 1981 favours).

If this second tradition is right, then the process is not, strictly speaking, an *English* Initial Fricative Voicing in that it would have been innovated either in a parent or sister language. To a large extent for our purposes here, it does not matter which of the traditions is correct. I shall argue below that the fricatives had fundamentally the same phonological identity in Proto-Germanic and up to the point when the EIFV occurred. While the ‘contact’ hypothesis is less easy to accommodate (and less likely to be true in my opinion), if the process occurred in Germanic, but only came to light when it was written in Middle English texts, or in another Germanic dialect and was borrowed, or if it was actually innovated around the time that it becomes visible in texts, the same thing happened. Whichever tradition reflects the truth, I shall retain the label ‘EIFV’.

As Nielsen (1981, especially chapter 2) reminds us, we should by no means ignore the possibility of polygenetic development for similar-seeming processes, even when they are relatively alike and in genetically related languages. Fisiak (1984) claims that there is no compelling evidence that the EIFV was innovated in common or West Germanic, writing that “it is at least equally likely that the development of initial voiced stops was a parallel independent development” (1984, 6). For this to be true, we have to be able to show that it is a relatively minor, plausible process, which we might almost expect to be innovated in languages with similar phonologies (or if not *expect to be innovated*, then at least we would not be surprised by it if it occurs). When we return to the EIFV in section 4.3, I propose just such an account.

2.1.3.4 Is the EIFV an example of lenition?

If we consider the type of process that the EIFV is an example of, whenever it may actually have occurred, it seems to be a clear cut case of lenition, given the simple definition adopted in section 1.4. This can be seen from the processes which are repeated here as (2.21). It is interesting to note in this regard that Fisiak (1984), in his detailed description of the EIFV,

writes “[t]here is no convincing evidence that there were four separate processes. On the contrary [...] it would seem that a unitary interpretation is the most satisfactory one. The voicing is in each case a phonetic process of weakening (lenition)” (1984, 4).

(2.21) EIFV

$f > v$

$\beta > \delta$

$s > z$

$\int > 3$

The place of articulation is not affected in any of these changes and they seem to be quite straightforwardly characterisable as a doubtless unitary case of lenition of the ‘sonorising’ type voiceless \Rightarrow voiced. We return to this process at various stages of the discussion below, and we will see in section 4.1.2 that the best way to interpret this process is quite unlike the traditional understanding.

2.1.4 Inner-German Consonant Weakening

The last of the four main pieces of data which we will discuss is well-known among Germanist philologists, but perhaps not elsewhere. While it clearly occurred centuries previously, and scholars had noted the effects, Lessiak (1933) is typically credited with being the first to recognise its processhood and having described and named it. He called it the *binnenhochdeutsche Konsonantenschwächung* which I translate as ‘the Inner-German Consonant Weakening’.²⁰ Kranzmeyer (1956) is also recognised to have contributed to our

²⁰ There does not seem to be a standard English translation for the title, which is doubtless partly a consequence of the fact that the process is not well-known outside of Germanist circles, where the German original can freely be used (as, for example, in Keller 1978 and Barbour & Stevenson 1990, following Schirmunski 1962). There is no generally recognised linguistic construct ‘Inner-German’ but there is no real construct ‘Binnenhochdeutsch’ either. The name reflects the geographical extent of the process which has not affected the peripheral High German varieties but has affected the inner-High German varieties. It would be problematic to translate *Binnen-* as ‘Central’ because ‘Central German’ is the usual translation for *Mitteldeutsch*, which is a technical term in German dialectology (see, for example Russ 1990). ‘Middle German’ would not be a good choice either as ‘Middle’ is generally reserved for a description of time rather than place (as in ‘Middle English’, ‘Middle High German’). I have not called it ‘the Inner-German Consonant Lenition’, which would be in line with Russ (1982), for example, because that would be to beg the question. I use ‘German’, rather than ‘High German’ partly because the process is also referred to as the *binnendeutsche Konsonantenschwächung* (for example, by Schieb 1970), but also because it did not affect all High German dialects, particularly if Kranzmeyer’s ‘Central Bavarian Consonant Weakening’ is extracted from it. The process has even been called the ‘Third Sound Shift’, for example by Keller (1978) and in part, in the curious Zabrocki (1965). This ties it in with the phenomena described here as the CGS and the HGCS in some uncertain sense, but I avoid this description here.

understanding of the phenomena which have been grouped together under the label, partly by splitting off a 'Central Bavarian Consonant Weakening', which is somewhat different but nonetheless occurred at around the same time and has a good deal in common with the IGCW. In this section, I focus on the IGCW, but to the extent that the two were the same, the points can be generalised.

While the occurrences dealt with in this section are discussed less frequently than the other data sets considered in this chapter and elsewhere, they had, at their most extreme, a substantial effect on the phonological varieties into which they were innovated. The effects of the IGCW can be heard in many present-day non-standard traditional dialects, in a rather similar way to those of the EIFV, and they have also not made it into standard forms of the language. Again, there is variation in the dialects as to what degree the process affected the input segments and also as to which segments counted as the input. The process is often situated as an aspect of Middle High German and is described in Handbooks which deal with the period and also in 'Histories of German' such as Paul (1944), Schieb (1970), Keller (1978) and Schmidt (1984). The effects of the process have frequently been noted in German dialect monographs, such as Albrecht (1881), Schübel (1955), Bock (1965) and Weldner (1991), and are described in general work on the relevant dialects, such as Bergmann (1991).

2.1.4.1 The IGCW: take 1

As with the HGCS and the EIFV, the IGCW was not innovated in exactly the same way in all dialects, but it was typically quite general and took as its input the stops which have been described above as *Ts*. The IGCW was innovated in dialects which had previously been affected by the HGCS₁, but not in the most southerly areas which show clear evidence for the HGCS₂. In terms of the input to the IGCW, this means that only those Germanic *Ts* which were not affected by HGCS₁ (those in environments where the process was inhibited) and those which had been borrowed in loanwords (from other languages or from German dialects which had not taken part in the HGCS) could form the input. In some dialects this added up to a substantial number of segments, especially in the areas where the HGCS₁ had been substantially inhibited. The inherited input then, in fact, derives at least in part from the GCS₃.

The processes given in (2.22) are a simple representation of the IGCW at its fullest extent, as was the case in the above exposition of the HGCS and EIFV. The processes formulated here are divorced from the phonological context of the system into which they were innovated

and are based on the correspondences given in Albrecht (1881) and Bergmann (1991) for the city of Leipzig specifically, (the same correspondences are described more generally by Schmidt & Vennemann 1985 for the 'Upper Saxon' variety, and Bergmann 1991 pinpoints the area as the South Osterländisch variety of Upper Saxon). The symbols are those of Albrecht (1881) and Bergmann (1991), and almost those of Weldner (1991), who describes a neighbouring Thuringian variety where the process has not consistently occurred in /k/ (in word-initial environment before vowels):

(2.22) IGCW

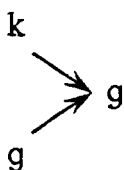
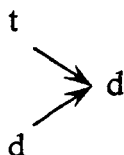
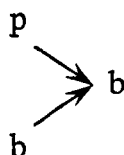
p > b

t > d

k > g

The dialect monographs typically describe the process as having occurred across the board. König (1978, 148) describes the IGCW as being carried out in "allen Positionen" [= 'all phonological environments'] for a broad swathe of Central and Upper German dialects. One point which has been glossed over in the presentation of the other processes in this chapter (either because it was irrelevant or because the details are uncertain) is how the innovation of a phonological process affects the rest of the phonology of the language into which it is innovated. The facts are clear for the IGCW. There were two series of stops before the process was innovated and, after the process had been lexicalised (because it was entirely context-free this could happen quite quickly), there was only one. The wider picture for the IGCW is, in fact, a set of segmental mergers as in (2.23):

(2.23) IGCW



Descriptions of modern dialects are often explicit in advocating IPA transcriptions which use one symbol such as [b] for both of the synchronic reflexes of what was an underlying contrast in stops before the innovation of the IGCW.²¹

When set out as in (2.23), this process, which might initially seem quite innocuous, is actually a counterexample to the generalisation discussed in section 1.2.2 which Hyman (1975) sets out, influenced by Jakobson and Greenberg, stating that no process could result in the voicing of all voiceless stops because it would leave an “impossible” phonological system because “/b, d, g/ implies /p, t, k/” (Hyman 1975, 17). The data is arguably therefore quite important for phonological theory. We will see later in this thesis, however, that the processes presented here are compatible with phonological tradition if certain novel analyses are accepted and that what is now Hyman’s (and Jakobson’s) paradox only *seems* to be a problem for phonological theory.

²¹ It is true, of course, that there are some other reflexes of these stops in these varieties due to processes which either preceded the IGCW and hence removed possible input, or followed it and affected the output, but this does not alter the point made here. We return to the discussion of stops featured in these dialect monographs in section 3.1.3.1. In certain present-day dialects a slightly different situation has led to the same output effect. Russ (1982) explains: “the reflexes of MHG /k/ and /g/ are kept distinct in initial pre-vocalic position”, not because they are maintained as stops, but because “MHG /k/ has become [g], while pre-vocalic /g/ has become [j]” (Russ 1982, 25). Although this is not Russ’s conclusion, there does not seem to be much of a case for a synchronic underlying contrast between /k/ and /g/ where this has happened, because there are hardly any realisations of MHG /g/ as [g] which would lead the child acquiring the language to posit such an underlying contrast in stops. It is doubtless true that in some varieties a minimal contrast in velar stops is maintained, however. Bergmann (1991) explicitly endorses the formulation given here, though; as well as the common merger in labial and coronal segments, he writes: “[i]n south Osterländisch NHG /k/ and /g/ have merged in a single phoneme /g/” (1991, 295).

2.1.4.2 Evidence for the IGCW

The type of orthographic evidence that can be expected following the kind of segmental merger outlined above is quite straightforward. If a contrast is lost which was once indicated in spelling by the use of two letters, one for each segment, then after the merger, we might expect to find either random alternation in post-process spelling or the consistent use of only one letter for both. The problem here, as for the EIFV, is that the process has been innovated in varieties which did not go on to form the standard and which have thus not been written a great deal. There is historical orthographic evidence for the IGCW, and precisely this kind of effect is shown in work on the errors made in the learning of standard German spelling (which continues the contrast between <p, t, k> and <b, d, g>) by children who are native speakers of modern versions of the dialects where the INGW was innovated (see, for example, Ammon & Loewer 1977, Zehetner 1977 and Kraemer 1978).

The best evidence for the IGCW, however, can be gained through the comparison of present-day dialect forms which exhibit the effects of the process with something which does not, in a similar fashion to the way in which the data from the *Survey of English Dialects* was used as evidence for the EIFV. The evidence in (2.24) compares Standard German (orthographic) forms, which indicate the historical and contemporary standard segment (which is shown in the first column), with forms from a range of non-standard dialects which show the effects of the IGCW. The non-standard forms are taken from Albrecht (1881), Schübel (1955), Bock (1965) and Weldner (1991); they are all intended as a form of phonetic transcription and I reproduce them as they appear in the sources:

(2.24) IGCW

p	Barchfeld German	<i>baged</i>	'package'	Waldau German	<i>buba</i>	'doll'
	Standard German	<i>Paket</i>		Standard German	<i>Puppe</i>	
t	Waldau German	<i>dif</i>	'deep'	Stadtsteinach German	<i>doxde</i>	'daughter'
	Standard German	<i>Tief</i>		Standard German	<i>Tochter</i>	
k	Leipzig German	<i>Gunst</i>	'art'	Barchfeld German	<i>grax</i>	'noise'
	Standard German	<i>Kunst</i>		Standard German	<i>Krach</i>	

2.1.4.3 Where and when did the IGCW occur?

This section has already dealt with certain facts about the dating of the IGCW, namely that it was innovated after the HGCS and that it occurred in slightly differing ways in different dialects. We are, unsurprisingly, again unable to be precise about the date of its innovation; Russ (1982) mentions that the dates which have been proposed vary from during the OHG period (ie, since the 8th century) to 1300 CE. There are clear orthoepic comments from the eighteenth century which show that the process must have been innovated before then. As Russ (1982) reports, Lessiak (1933) does not date the process, but does identify the main IGCW with the dialect areas of Low Alemannic and all of Central German, except Silesian, North Thuringian, Riparian and Moselle Franconian (this excludes Kranzmeyer's (1956) 'Central Bavarian Consonant Weakening'). It seems reasonable to assume a date relatively early in the second millennium CE.

2.1.4.4 Is the IGCW an example of lenition?

There seems little doubt that many authors regard the IGCW as an example of lenition. The very name given to it by Lessiak (1933) claims as much, because *Schwächung* is the non-Latinate loanword German equivalent of *Lenierung*, just as *weakening* in English is for *lenition*. Some German authors give both descriptions.²² Keller (1978, 277) overtly describes the process as a lenition and indeed, the simple form of the process, repeated here as (2.25), fits onto the lenition trajectories given in section 1.4.

(2.25) IGCW

p > b

t > d

k > g

Like the EIFV, this seems to be a quite straightforward case of lenition of the 'sonorising' type voiceless \Rightarrow voiced. As we will see when we return to this process in section 4.1.4, this apparent straightforwardness is somewhat misleading.

²² For example, Paul (1944): "[es] folgte auch der zweiten Verschiebung eine Schwächung (Lenierung), die Teilweise weit zurückreicht" (1944, 74).

2.2 Lenition and Germanic: summary and prospect

The four sets of data examined in this chapter all represent quite substantial changes in the phonology of the languages that they were innovated in. They are not entirely exceptional, though, and entirely equivalent processes could be found from other language families, and indeed from within the Germanic group of languages, too.²³ A fair amount of the discussion in the later chapters of this thesis will derive from a careful consideration of these processes, although first, in section 2.3, a slight case of simple reinterpretation will be visited on some of the processes.

The provenance of the four sets of processes ensures the 'Germanic' bent of the thesis, but it could be seen as problematic. One concern might be whether the data discussed here constitutes a wide enough database for generalisation because they are all related languages and so might not form a firm basis for universalist phonological considerations. I do not think this should give us any real cause for typological worry, however. It is true that I am chiefly considering 'genetically related' languages (in that they are all typically described as being descended from Proto-Germanic) but there are several points to note here.

One simple point in defence of the approach taken in the thesis is that section 3.2 brings in a range of data from other languages, some of which are more distantly related (deriving from other branches of Indo-European) and some of which are not generally thought to be related to Germanic at all, given our present state of knowledge of genetic relationships between languages.²⁴

Another point in defence of the position taken here is that it may be that observations and generalisations set up on the basis of the investigation of one type of language (and it is possible, although not that straightforward, to consider a language 'family' as a 'type of language') should only apply to that type of language. This does not seem to be the null

²³ One major Germanic lenition that I do not consider here is the substantial set of processes that have occurred to stops in Scandinavian languages (see, for example, Haugen 1976). These do not feature here for reasons of space, but are compatible with what is finally proposed. It would take a substantial discussion to deal with the way in which the processes have spread differentially in the dialects of the Scandinavian languages. They are at least partly similar to the Spanish lenitions which are briefly discussed in section 3.2.

²⁴ This is not the place to consider so-called 'long-distance' historico-comparative reconstruction, in which several researchers have loosened the comparative method in the search for possible word-wide cognates (Ruhlen 1987, for example, is explicit in his rejection of the comparative methodology). This kind of work has not met with much acceptance in the historical linguistics community. Salmons (1992), Fox (1995), Lass (1997) carefully weigh up this newer methodology and some of the claimed results and point out the flaws in historical reasoning which become possible if the comparative method is rejected.

hypothesis to me, but even if it were, it could still be argued that this is compatible with the approach taken in this thesis. The implications of the study of certain effects in Germanic languages are extended to account, at least in part, for other facts from Germanic languages. I would be reluctant to withdraw to this 'language-family-specific' position, however, and given the consideration of data from other languages in section 3.2, I do not feel it necessary.

A further point is that any sample of languages is inescapably not going to be representative of all possible human languages. It is surely the case that most possible human languages have never been spoken and many more have disappeared without being recorded. We know the history of fewer still. At times, it is wise to work with data which we can be relatively sure of. In any case, if we take the notion of 'universalism' seriously, and ignore the ideas that related languages 'drift' in the same way or the mystical idea that there is a 'spirit of the language' guiding it in some direction, then, as long as languages are separate from each other (in time or space) and not in contact, we can view events as different and separate developments, even if they occur in genetically related languages.

The Germanic bent is thus indeed intended, but should not impinge on general theorising. There is certainly much to consider: it seems that the ICCW is a case of an 'impossible' merger from voiceless sounds to their voiced congeners. The EIFV seems to be a similar type of process, but it did not involve a merger because there was only one set of underlying fricatives in the phonological system of the language when it was innovated. The GCS and the HGCS both involve several quite possibly unconnected processes, some of which show intriguing patterns of exceptions. The reinterpretations to come will be quite substantial in part, but we will approach them gradually. The first stage in this is to bridge the gap between philology and linguistics.

2.3 Realism in reconstruction in historical phonology

At times, reading Handbooks of the type which were used to construct section 2.1 can give the impression that the diachronic phenomena that they deal with are unique phenomena which are not connectable with other processes in other languages. It may well be that the act of naming them so forcefully (*the High German Consonant Shift*, for example) can contribute to this impression and, although it was probably not normally the intention of the original authors, the effect can obtain, nonetheless. Of course, the Germanic developments discussed in section 2.1 cannot be different in type from what has happened in the diachrony other languages and

from what is generally possible in language change. If we are to banish this possibility of Germanic insularity, it is important to view the processes in a language-universal perspective. Firstly, it is possible that analogous events in other languages at other times might help us to work out the best way to interpret exactly what happened in the Germanic developments just discussed, and secondly, it may be that simply reformulating the processes using a more universalist vocabulary and symbolism will facilitate our understanding.

This section does both these things and a little more. In general, it confronts the changes as they were just portrayed with various types of more general concerns. For some of the changes, this will lead to a slight reformulation. Where this is so, it is mostly not a matter of substantially changing the data, but rather of re-representing it so that it fits with contemporary ideas. In one case, however, a quite major change is brought to the presentation, which is nonetheless relatively straightforward to digest, and, once discussed, will cause us no further concern. In another case, a little more uncertainty will set in. This will not fundamentally alter the vital data for this thesis but it will introduce some important thinking and doubts. It would be false to suggest that the formulations of the processes as given in section 2.1 are those accepted by all philologists or historical linguists, but the formulations in that section are accepted by many and are 'standard' positions. This section functions to bridge the gap between philology and linguistics by linking the data with universalist phonological concerns. This involves showing both how others have already applied such ideas to reinterpret the data and also by bringing in some novel argumentation.

The factors considered in this section are actually quite diverse and do not have a simple thematic unity but they all have the potential to alter the simple symbol to symbol equations that were given as summaries of the four sets of processes in section 2.1. They can be seen as focusing on the desire to ensure that the reconstruction of past phonological states and processes (which is a shared aim of philology and historical linguistics) should be as realistic as possible.

The first subpart of this section (2.3.1) undertakes some relatively simplistic reinterpretation (and some that is not so simplistic) in that it 'rationalises' the formalisation of the processes from section 2.1 into the transcription conventions of the International Phonetic Association's Alphabet (the IPA). The second and third sections (2.3.2 and 2.3.3) both focus on issues which are largely relevant to the GCS, but will also have a certain impact on the analyses that are to be proposed for the other processes later. They deal with the two main aspects of what has been called 'the new sound of Indo-European' (in Vennemann 1989a, for

example), the Laryngeal Theory and the Glottalic theory. We will see that they have different kinds of implications for the data and some of the discussion there will be familiar to those acquainted with the debates, as it is by now classic argumentation. However, some of it (eg, the Glottalic Theory) cannot be ignored, and the rest will have intriguing and perhaps unexpected resonances elsewhere in the thesis.

The question of how 'real' the reconstructions of the phonology of dead languages can be (and can be expected to be) is a live issue in historical phonology and it is a natural question to ask, especially in connection with those phonological systems which have been proposed purely on the basis of the comparative method of reconstruction. This is highly relevant to the reconstruction of IE and, as Fox (1995) notes, the discussion in this area has generally focused on the phonological system of that language, but the issues are just as relevant to any other past phonological state, such as, for example, that of Germanic. The segments that I have, perhaps foolhardily, transcribed without the 'reconstructed' asterisk in this thesis are, of course, not ones for which we have direct phonetic evidence. There has been quite some debate as to (i) their ontological status (ie, should we view them as simply useful cover symbols for correspondence classes or should we interpret them as representing phonological segments that were once as psychologically real for speakers as a transcription like /f/ is interpreted for Present-Day English) and (ii) if we take the latter option, what precise phonological value should they be assigned.

These issues have been addressed recently in such works as Vennemann (1989a), Salmons (1993), Schwink (1994) and Fisiak (1997) (and see also the discussion in Fox 1995 and Lass 1993, 1997), but they have a long history. In his discussion of the issues, Koerner (1989) describes how the key two positions on the question might be identified with Schleicher and Saussure. Schleicher (for example 1868) was optimistic that the best methods of philology could reconstruct the true system for IE while Saussure regarded reconstructed segments as quasi-mathematical values in an abstract system. In terms of recent thinking, Salmons (1993) identifies a "current shift away from abstraction in reconstruction and toward increased attention to phonetic detail" (1993, 72), which seems reasonable, given that, in the words of (Lass 1997, 272) "a reconstructed language is a natural language, just like any other". We expect the uniformitarian principle to apply and, as argued at the start of this section, we can only hope to truly understand the phenomena of historical phonology if we are able to compare them to each other and also to the phenomena of the present-day synchronic states of

languages. We can only do this if a symbol like ‘g’ means the same in all reconstructed phonologies and has the same ontology there as it does in the description of synchronic phonology. I assume in what follows that we can and that it does.

The issues of phonological symbols brings us to the topic of the next section, where I reinterpret some of the data from section 2.1 to make sure that the symbols used do have an easily comparable form.

2.3.1 The use of symbols and their implications

As I discussed at the time, the presentation in section 2.1 used the symbols that are found in the Handbooks, which are the repository of philological achievement. The symbols are mostly easily interpretable for the phonologist, but are not always those used in modern historical linguistics. It is entirely straightforward to remedy some of this but there are a few issues that will require further discussion in this procedure. In this section I replace the symbols that are not part of the IPA with ones that are. The IPA has been constructed with universal applicability as its primary concern (see IPA 1999 for details) and is standardly assumed in much phonological writing.²⁵ The IPA is not constructed entirely on phonological grounds (for example, there are no unitary symbols for affricates, see Durand 2000) but it is indisputably the most widely applicable and widely applied system.

The symbols given in (2.26) are easily replaced. Some of the philological conventions, such as the use of thorn for the voiceless dental fricative, are found in relatively direct forerunners of the IPA, such as Sweet (1908) and their value is clear. As well as thorn, the symbols for voiced fricatives are also easily replaced, as shown in (2.26).

(2.26)

þ = θ

ḥ = β

ḍ = ð

ḡ = γ

²⁵ It is given as a preliminary, assumed beginning in phonology textbooks from Britain, at least (eg, Carr 1993, Spencer 1996, Davenport & Hannahs 1998, Roca & Johnson 1999). The American tradition is to use a slightly different set of symbols, but here the differences really are minimal.

It is also straightforward to IPA-ify the *TA* and *MA* symbols in IPA terms: it would be normal to raise the following aitch to a superscript, given that the symbols are interpreted to represent unitary segments. This is shown in (2.27) and the segments are thus indicated to feature underlying aspiration, as is conventional in the transcription of certain segments of Hindi, for example:²⁶

(2.27)

$$\begin{aligned}bh \ dh \ gh &= b^h \ d^h \ g^h \\ph \ th \ kh &= p^h \ t^h \ k^h\end{aligned}$$

Some of the other non-IPA symbols used in section 2.1 are not so simply replaced, however, although for different reasons. There are two such cases. The first concerns the symbols used for the dorsal segments which are output to GCS_{1a} and GSC_{1b}, and for the dorsal output of the HGCS₁. The GCS processes are repeated here, with newly superscripted aitches, as in (2.28).

(2.28)

$$\begin{aligned}k^h &> \chi \\k^{wh} &> \chi^w \\k &> \chi \\k^w &> \chi^w\end{aligned}$$

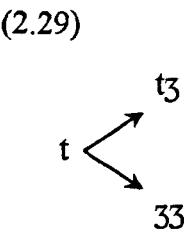
The situation is the same for all four correspondences (and for the HGCS) so they can be treated as one. If these were IPA symbols, the input would be velar and the output uvular but this is not how the correspondences are generally understood. Picard (1999), for example, makes a point of explaining that the chi symbol is intended to be understood as a velar. This would make the process a simple spirantisation with no change of place at all. However, it is not entirely clear that this is the correct interpretation of events. The symbol /x/ unambiguously stands for the voiceless velar fricative in the IPA and it is often used to transcribe segments occurring in German and Dutch, for example, some of which are the direct descendants of the Proto-Germanic fricatives, which were derived in the GCS, others (in German) were derived in the HGCS. The problem is that in the Present-Day languages, the segments are typically not realised as velars but are indeed uvular (they are sometimes described as pre-uvular or post-

²⁶ This demands more attention, though, and will receive it in the next two sections.

velar). Kohler (1977, 1990) and Scheer (1998) are explicit in this regard for German, as are Collins & Mees (1981) for Dutch.

What is more, there are other cases where the IPA symbol /x/ has been used to transcribe segments which are typically uvular, for example in Scots, where the fricatives are also often the output of the GCS (see Stuart-Smith 1999). In addition, voiceless fricatives in Liverpool English, which are derived from underlying /k/ (see Wells 1982, Honeybone 2000c, 2001) are typically uvular. This might well lead us to conclude, with Scheer (1998), that the output of the GCS was uvular, rather than velar. We cannot engage with all the issues here, however, and timidly retreat to the use of IPA /x/ for the output of these processes. There is reason to doubt that truly velar fricatives occur, at least underlyingly and at least in Germanic; this has consistently been disguised by the use of IPA /x/ to transcribe them and I maintain that tradition here. It may well simply be that the natural lenition output for /k/ is not truly velar. In fact, however, similar issues arise for other segments, although none of the processes involve a change in the major articulator involved.

The second case where it is not so straightforward to replace a philological symbol with an IPA equivalent is the case of ‘3’. This was used for outputs of the HGCS₁; the relevant part of process is repeated here as (2.29):



Although there is a segmental split here at some level, the point is the same and we can largely treat both outputs together by referring to the segment as ‘3’, ignoring the issues of geminacy and affrication for the moment (these are dealt with in detail in chapter 4). The issue here is: what is the precise phonological nature of the fricative that this symbol represents? This has long been clearly recognised as an issue and all Handbook discussions of the HGCS devote some space to it. Several things are clear. There was a sibilant in Germanic which German has inherited, (for example in such words as *küssen* ‘kiss’ and *Eis* ‘ice’). This sibilant and the fricative product of the HGCS have merged, so that in the phonology of Present-Day German there is no evidence that they were ever distinct (Paul 1916 assumes that this merger happened

in the second half of the thirteenth century²⁷). Equally certain is the fact that this merger did not occur immediately after the innovation of the HGCS. As Joos (1952) explains “[o]ver a century ago, rime studies firmly established the fact that High German, for half a dozen centuries at least down through the classical Middle High German period (MHG, ca. 1200 A.D.), had two voiceless spirant phonemes, both sibilant as far as we can tell...” (1952, 373).

The philological evidence is clear: even if the segment is written <z> and not <ʒ>, this means that it is not distinguished from the affricate; it did not merge orthographically with the Germanic <s>. The rhyme evidence that Joos refers to (for example, that words such as *wizzen* ‘know’ and *wissen* ‘knew’ were not used as rhymes in poetry) is conclusive. The fact that the two did not merge will be important later.

There have been many proposals for the realist interpretation of ‘ʒ’ (which often consider the precise realisation of the Germanic sibilant, too). Joos (1952) assumes that ‘ʒ’ was an apical fricative and that the inherited Germanic fricative was pre-dorsal. Russ (1982) lists several others: for example, Penzl (1970) assumes that ‘s’ was retroflex, whereas ‘ʒ’ was non-retroflex; Keller (1978) assumes that ‘ʒ’ was a dental fricative, whereas the other was more palatal.

It seems that there are real problems in ascertaining the precise phonological and phonetic identity of the segment which has been transcribed ‘ʒ’ (and the other sibilant, ‘s’ for that matter). It may also seem difficult to see what kind of evidence could help settle the matter. One possible source of such evidence would be to compare the HGCS₁ output with the output of an analogous process in another language; to bring in the ‘universalist’ idea discussed in the introduction of this section. If historical processes are natural, then we expect to find the same process in more than one language, and where one such process analogue is clearly observable, this could offer important evidence for the reconstruction of past processes.

Luckily, there is such evidence, although it has not previously been recognised. A remarkable analogue to the HGCS₁ is found in the variety of English which is spoken in and around the city of Liverpool, England. The process is synchronic and variable (as might be expected in a process which clearly marks out a variety as non-standard). It is a generalised lenition process (on the definition given in section 1.4) which has the stops /t, k/ and to a lesser extent /p/ as its input and has voiceless affricates and fricatives as its output. The details of

²⁷ “Ein Zusammenfall von *s* und *ʒ* erfolgte in der zweiten Hälfte des 13. Jahrh.” (Paul 1916, 343).

where affricates can occur and where fricatives occur are strikingly similar to those of the HGCS₁. The precise details of the process are yet to be fully described²⁸ but preliminary findings are reported in section 3.2.1.8 of this thesis (and see Knowles 1974, Wells 1982, Sangster 1999, to appear, and Honeybone 2000, 2001, for discussion of aspects of the phenomenon). We can be quite sure that the process was introduced into the variety in the nineteenth century (see Knowles 1973) and it is the best analogue for the HGCS which has up till now been reported.²⁹

Further details can wait until section 3.2.1.8, for what is vital here is the precise nature of the fricative and affricate output for /t/. Just as in the post-HGCS situation in German, the fricatives are not identical with realisations of the pre-existing /s/ (Sangster 1999 shows this instrumentally), which is even more encouraging as to its analoguehood to the HGCS₁. The discussion in section 4.1.2 of this thesis will make it even more analogous by dealing with questions of geminacy.

The fricative output is a slit fricative, lacking the tongue groove which is typically associated with a classical alveolar fricative and is best compared to the ‘slit-t’ described for Hiberno-English in Pandeli, Eska, Ball & Rahilly (1997), following key work by Hickey (1984, 1996, 1999). The fricative is a “controlled articulation” (Hickey 1984, 234) but there is less contact with the roof of the mouth than for canonical [s] as it has a broad central channel, as Pandeli et al (1997) show through electropalatography. Pandeli et al (1997) note that there is no unitary symbol for the segment in the standard IPA and, after discussing a range of possible transcriptions, suggest [θ̥] which is composed of the base symbol [θ] which clearly indicates fricativity and a flat cross-sectional tongue shape and the diacritic [̥] which indicates a precisely alveolar place of articulation. It is taken from the ‘extended IPA’ which is used to transcribe disordered speech (see Duckworth, Allen, Hardcastle & Ball 1990). Following Pandeli et al’s (1997) suggestion for what seems to me an very similar segment, I transcribe the Liverpool English lenition-derived fricative and (secondary portion of) the affricate as [θ̥] and [tθ̥].

²⁸ Research on these issues is in hand at Edge Hill College.

²⁹ The similarity between the HGCS₁ and the situation in Liverpool English has been noted before by scholars who are acquainted with the diachronic phonology of German, for example in Hock & Joseph (1996), and Knowles (1974) reports on an observation to this effect by Kuno Meyer, who was lecturer in German (and subsequently professor of Celtic) at Liverpool University. He had trained in Leipzig under the Neogrammarians (see Ó Lúing 1991). The parallelism between Liverpool English and the HGCS has never been properly investigated, however.

It is interesting to note that the typical tongue position of the underlying ‘alveolar’ fricative in Liverpool English (ie, that which would typically be transcribed as /s/) involves “just the blade, with the tip down by the lower teeth” (Knowles 1994, 106)³⁰ which is similar to the equivalent segment in Present-Day Dutch (see Collins & Mees 1981). The Dutch segment, which is etymologically the same as OHG ‘s’ could well be a modern reflection of a realisation that was once shared with German. This might allow us to draw conclusions about the precise nature of OHG ‘s’.

In fact, the whole scenario is rather similar to Keller’s (1978) and others’ suggestions for the distinction between OHG ‘z’ and ‘s’, and the data from the HGCS-analogue in Liverpool English could be the only piece of reliable phonetic evidence for the OHG distinction. I propose, therefore, that the realist interpretation of the segment ‘z’ should be the slit alveolar fricative [θ]. This gives us the equivalences in (2.30):

(2.30)

t_z = t_θ

z_z = θθ

This section has seen the reformulation of several of the processes that were described in section 2.1. I re-summarise the processes here, to conclude this section before we move on to other issues. The IPA-ification of the processes has affected three of the processes, the GCS, the HGCS and the EIFV. I give their new forms below as (2.31), (2.32) and (2.33) respectively. The IGCW has not been affected by this reformulation.

³⁰ Knowles (1974) is the standard reference work for Liverpool English and includes a wealth of important data, (although the author does not engage in extensive phonological analysis). Knowles’ description of the fricative realisation of /t/ is unusually undetailed, however; he describes it as “incomplete /t/” (1974, 327).

(2.31) The GCS: **take 2**

GCS_{1a}

$$p^h > f$$

$$t^h > \theta$$

$$k^h > x$$

$$k^{wh} > x^w$$

GCS_{1b}

$$p > f$$

$$t > \theta$$

$$k > x$$

$$k^w > x^w$$

GCS₂

$$b^h > \beta$$

$$d^h > \delta$$

$$g^h > \gamma$$

$$g^{wh} > \gamma^w$$

GCS₃ [not altered but included for completeness]

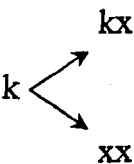
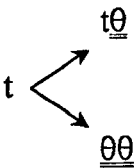
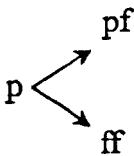
$$b > p$$

$$d > t$$

$$g > k$$

$$g^w > k^w$$

(2.32) The HCGS: **take 2**



HGCS₂ [not altered but included for completeness]

b > p

d > t

g > k

(2.33) The EIFV: **take 2**

f > v

θ > ð

s > z

ʃ > ʒ

2.3.2 The Laryngeal theory

After the relatively straightforward reformulation of the last section, we turn now to a more radical revision. The revision relies on some of the most insightful proto-phonological thinking ever in historical phonology and is a case study in how reconstructive hypotheses can find important corroboration. The material discussed in this section is clearly not intended to be thought of as a new contribution to discussion of the processes, but the presentation will be important for coming points which will be more novel.

The issue has come to be known as the ‘Laryngeal Theory’ of Indo-European consonantism (see, for example, Polomé 1965, Fox 1995).³¹ As is often discussed, Saussure (1879) proposed that IE had a set of segments which he labelled ‘coefficients’. There was no simple evidence for these segments in that they were not written as letters in any of the attested IE languages that were known about at the time, and Saussure’s brilliant leap in methodology (described in detail in Fox 1995 and elsewhere, as is the rest of the story) was to reconstruct the segments using the methods of internal reconstruction, ie, original segments were projected back from the phonological effects that they could be seen to have had on other segments, even though they had been lost as segments themselves. The discovery of Hittite provided some remarkable

³¹ As Vennemann (1989b) points out, comparing the Laryngeal Theory and the Glottalic Theory, which is to be considered in the next section, “[l]ooked at from the point of view of the philosophy of science, the grand name of Theory is not really warranted for either of the two approaches. They only form portions of theories, such as the phonological, morphological, and lexicological theories of Proto-Indo-European and theories of the phonological development of the Indo-European languages” (1989b, x). However, I retain the traditional terms. The same point about the use of ‘theory’ might well hold for certain other occasions where the term is used in this thesis, but I go with popular usage.

confirmation of Saussure's proposals in the hands of Kuryłowicz (1927), thanks to the types of segments which were written in the cuneiform in environments that corresponded with those predicted by Saussure. This has widely been taken as evidence for the existence of Saussure's coefficients. Saussure was not particularly concerned with the realist interpretation of the segments that he proposed (although he transcribed them as being vocalic). Subsequent research has shown them to be consonants, more specifically 'back' consonants, hence the name 'laryngeals'.

There has been considerable discussion as to how many laryngeals there were and how they should be interpreted realistically. It would go beyond the scope of this section to consider all the proposals that have been made in this regard, even though this would fit with the 'realist' theme of the supersection (see many of the papers in Vennemann 1989a for some recent discussion of this topic). For our purposes here, it is simplest to follow Bammesberger (1989) and, as Bammesberger reports, Kuryłowicz's own later ideas, that there was only one laryngeal.

The key relevance of this to our purposes here is that we can be quite sure that laryngeal segments existed in IE. This has had an important effect on the reconstruction that is generally accepted for the system of IE stops. The system given in (2.1) had four series, but there was never substantial evidence for the *TAs* (/p^h, t^h, k^h, k^{wh}/). This is in fact shown in the 'evidence' section of 2.1.1 where no examples could be adduced for the changes in the putative segments /k^h, k^{wh}/ . The main evidence for their existence is from the fact that Indo-Iranian had four series of stops, as do some of its modern daughter languages. Since the advent of the Laryngeal Theory, however, it has been discovered that the Indo-Iranian *TAs* were principally derived in that language from a cluster of a *T* followed by a laryngeal (see Lehmann 1993, Fox 1995, Beekes 1995 for details).³² This means that a process was innovated in Indo-Iranian which had the effects shown in (2.34), where /h/ stands for a laryngeal.³³

³² The small amount of evidence for IE *TAs* from Greek and some of the other Indo-Iranian *TAs* has been accounted for in other ways.

³³ To an extent, the philological transcription that was abandoned in section 2.3.1 was thus not far wrong, in that it almost seems to indicate sequences of stops followed by laryngeals, but that was not how it was meant - *TAs* were originally considered to be underlying unitary segments in IE. Also, I have not included the labiovelar; the history of IE dorsal segments is complex and would require too much space to discuss here (see, for example, Lehmann 1993, Beekes 1995); nothing in this history affects the points made here.

(2.34) Indo-Iranian TA formation

ph > p^h

th > t^h

kh > k^h

We will return to this type of process in chapter, as we will for the other types of processes discussed in this chapter. The process in (2.34) is not thought to have been innovated in Germanic, nor in other IE languages; the laryngeals had other effects but were generally deleted, so no TAs were formed. It is not unusual to lose segments such as /h/ in languages.

The most frequently accepted reconstructions of IE now have only three series of stops.³⁴ This means that the correct formulation of the GCS is not, in fact, as was given in section (2.3.1) because there simply was no GCS_{1a}; all the segments that were covered under that are now subsumed under GCS_{1b}, which gives us a simple GCS₁, and the whole 'shift' is now as set out in (2.35):

(2.35) The GCS: **take 3**

GCS₁

p > f

t > θ

k > x

k^w > x^w

GCS₂ [unchanged from 'take 2']

b^h > β

d^h > ð

g^h > γ

g^{wh} > γ^w

GCS₃ [unchanged from 'take 2']

b > p

d > t

g > k

g^w > k^w

³⁴ In fact, this is often the case in the Germanic Handbooks, but not always.

This is not quite the last word on previous interpretations of the GCS, however, as we will see in the next section. This current section is not just intended to recapitulate the historiography of IE historical phonology. The processes involved will be important later.

2.3.3 The Glottalic Theory

This final section in this chapter returns again to the status of the stops in Indo-European, as the investigation of the Laryngeal Theory in section 2.3.2 has just done, but it follows a rather different sort of argumentation and the effect of the discussion will be much less clear. The story of IE consonantism in the twentieth century (as told by Lehmann 1993, Salmons 1993, Fox 1995, for example) starts off with the four series system, inherited from Brugmann & Delbrück (1886-1900) and assumed here in section 2.1. It then undergoes the Laryngeal Theory, to emerge with the three series of *Ts*, *MA*s and *Ms*, assumed here in section 2.3.2, a change which most scholars followed, although a few still maintain the four series system. It then undergoes the Glottalic Theory, to emerge in a rather uncertain state, with some researchers avidly proposing a further revision of the stops and others steadfastly refusing it. Even among those who propose a revision, there is disagreement as to exactly what the realist interpretation of the IE stop segments should be. In this section I briefly examine the arguments, which are so widely discussed in the IE and Germanic literature that they demand our attention, and attempt to move towards a solution that will allow us to progress.

The key problem that the Glottalic Theory seeks to solve is, in fact, a result of the Laryngeal Theory's success. Once the *TAs* were eliminated from the system, the remaining stops in IE were the *Ts*, *Ms* and *MA*s, which were normally glossed as 'voiceless', 'voiced' and 'voiced aspirated'. The problem is, as Lehmann (1993) puts it, that "[n]o-one with any background in phonology would assume a set consisting of voiceless stop, voiced stop, voiced aspirated stop, the dental set of which might be represented *t d dh*" (1993, 74). There is no clearly attested case of a language with precisely this system of laryngeal contrasts. The identification of the problem thus relies on the use of typology in linguistic reconstruction, which is a basic form of phonology, asking the question 'what is possible in the phonological system of languages?' Clearly, this is a vital point. As was explained in the introduction to the current section 2.3, once we assume the uniformitarian principle, then it is not acceptable to reconstruct a language which has what is, to the best of our knowledge, an impossible phonology.

The phonological problem with the IE stops has been recognised for at least half a century now, and the ‘Glottalic Theory’ is an attempt to solve the problem by bringing the IE system into line and to give it a system which is phonologically plausible. As Salmons (1993) explains “[i]t is customary to trace the inspiration for the Glottalic Theory back to Jakobson’s 1957 presentation (published in 1958 and again in 1971), with frequent mention of work by Pedersen (1951) and Martinet (1955).” He goes on to trace it back even further, as does Collinge (1985). The name ‘Glottalic Theory’ derives from certain reformulations that were proposed in the early 1970s which feature non-pulmonic stops, ie, those with laryngeal specifications other than ‘voiced’, ‘voiceless’ and ‘aspirated’, such as ejectives.³⁵ The name is often used to refer to a whole string of proposals which have been proposed for the reinterpretation of the IE stops since the 1970s, largely following the lead of those whose proposals included an ejective series, but not necessarily using ejectives themselves.

The early substantial and separately-formulated proposals are Emonds (1972), Gamkrelidze & Ivanov (1973) and Hopper (1973). Their proposals are summarised in the table in (2.36), which lists the traditional reconstructions for coronal stops against the symbols proposed by the authors:

(2.36)

Traditional	Emonds	Gamkrelidze & Ivanov	Hopper
t	t ^h	t ^(h)	t
d ^h	d ^h	d ^(h)	<u>d</u>
d	t	t'	t'

It can be seen from (2.36) that Emonds’ (1972) account does not, in fact, include ejectives in his reconstruction, but it still belongs with the ‘literal’ Glottalic Theory proposals because it proceeds from the same reasoning.³⁶ Gamkrelidze & Ivanov (1973) actually use the symbols /t^h/ and /d^h/, but their model of phonology is a little inconsistent. They propose that aspiration is allophonic, but it is not entirely clear which segment is intended to be underlying. They write: “[t]he feature aspiration is a redundant property of the phonemes in this kind of system.

³⁵ The proposal is also called ‘the ejective model’, for example in Job (1989).
³⁶ Although Emonds has not contributed further to the question, his paper did have an impact; it is cited by Hopper (1973), for example.

From a strictly *phonological* viewpoint, the three series could be described as glottalised/voiced/voiceless. The phonetic property of ‘aspiration’ is, however, a relevant feature of the relevant series of stops which explains their diachronic changes and their reflexes in the historically attested languages” (1973, 155).³⁷ From this, it seems that the unaspirated segment is viewed as underlying and as being accompanied by a contextual process of aspiration.

Hopper’s /d̤/ is described as ‘murmured’, following Ladefoged (1971). This is intended to represent the kind of stop still found in Indo-Aryan languages, which are also described as having ‘breathy voice’ (see Ladefoged & Maddieson 1996). It is also basically equated to the laryngeal specification of ‘voiced aspirated’ stops in Ladefoged & Maddieson (1996), so it is not clear that there is a difference and Hopper’s symbol may be seen as a notational variant of /d^h/.³⁸

There are reasons other than simple inventory typology for postulating a series of ejectives: it accounts better for the infrequency of the labial in this series and for certain segmental co-occurrence restrictions in roots (see Gamkrelidze & Ivanov 1973), so it is not conjured out of the air. Other proposals have been made but it would go way beyond the scope of this section to evaluate them all, or even several of them, in detail. It is probably the case that Gamkrelidze & Ivanov’s proposal has found most support among other researchers (for example Bomhard 1986 accepts it entirely and Vennemann 1984 accepts the ejective series at least). Gamkrelidze & Ivanov’s is certainly the best thought through, and the authors have produced Gamkrelidze & Ivanov (1984), which is a compendious modern IE Handbook, dealing with the implications of their reconstruction for an extensive range of IE linguistic phenomena.

As mentioned above, the Glottalic Theory has not met with unanimous acceptance. Some Indo-Europeanists have rejected the whole idea that the traditional system is in need of revision, while others have rejected the idea that one of the IE stop series was comprised of ejectives but have proposed other reconstructions which seek to remedy the typological

³⁷ The original reads “Das Merkmal Aspiration gilt in einem derartigen System als redundante Eigenschaft der entsprechenden Phoneme. Vom streng *phonologischen* Standpunkt aus könnte man die drei genannten Serien als glottalisiert/stimmhaft/stimmlos kennzeichnen. Die phonetische Eigenschaft «Aspiration» ist aber ein relevantes Merkmal der entsprechenden Serien der Verschlusslaute, das deren diachrone Veränderungen und die Reflexive in den historisch belegten Sprachen erklärt.”

³⁸ Ladefoged & Maddieson (1996) actually use /d^h/ with a superscript ‘voiced glottal fricative’ symbol, but this is probably also just a notational variant for ‘short burst of breathy voice’.

implausibility in different ways. One reasonable problem that has been noted with the Glottalic Theory is that none of the standard daughter languages have ejectives. This is not a killer argument (none of the Present-Day daughter languages have laryngeals, either, but they are widely accepted) but it is a fair reason why the proposal has not been accepted without question. Nonetheless, as Salmons (1993) concludes: “[w]hile there is some evidence to support the claim that the Glottalic Theory is becoming the standard view of [Indo-European] obstruents, it is somewhat clearer that the old system is dead. The attempts to counter the Glottalic Theory without exception propose other alternatives rather than defending the system canonized in the great works from Schleicher to Lehmann” (1993, 72).

It seems reasonable that some revision of the classic IE system is called for and this would have effects on our formulation of the GCS. It is not so simple to pick a version of the Glottalic Theory and assume that it is correct, though. The discussion of laryngeal specifications will also feature later in this thesis, where certain revisions are proposed, to the extent that it would not be helpful to simply accept or reject the Glottalic Theory here. I thus give an either/or version of the GCS in (2.37). This retains the IPA versions of the traditional symbols next to Gamkrelidze & Ivanov’s version of the Glottalic Theory, which, as mentioned above, was one of the first versions to be proposed and has been tested and explored the most:

(2.37) GCS: **take 4** (‘traditional’ *or* ‘Glottalic’)

GCS₁

{p *or* p^(h)} > f

{t *or* t^(h)} > θ

{k *or* k^(h)} > x

{k^w *or* k^{w(h)}} > x^w

GCS₂

{b^h *or* b^(h)} > β

{d^h *or* d^(h)} > ð

{g^h *or* g^(h)} > γ

{g^w *or* g^{w(h)}} > γ^w

GCS₃

{b *or* p’} > p

{d *or* t’} > t

{g *or* k’} > k

{g^w *or* k^{w’}} > k^w

This take on the GCS is intended to be realist, unlike, perhaps, ‘take 1’ in section 2.1.1. It is noncommittal at the moment as to which of the realist values are to be recognised as that which truly coincides with the past psychological reality of IE speakers, but the claim is that, in principle, one of them did. There are still certain problems with these interpretations of the shift, however, as we will see. This spring largely from the fact that the processes involved have not been constructed with the aid of modern phonology. We return to them in detail in section 4.1.1.

2.3.4 Summary and prospect

This chapter has introduced four sets of historical phonological data. The earlier sections presented a ‘standard Handbook-type’ view of the data from a chiefly philological perspective and showed that there are indisputably some phonological developments in each of the four cases that historical phonologists need to be able to account for. The latter sections have linked this data to certain more linguistic, phonological concerns and have shown that certain aspects of the earlier presentation require or invite reanalysis. The kind of reanalysis that they have been subjected to in section 2.3 has been, at times, substantial but nonetheless, has only been informed by sometimes simplistic phonological ideas, such as the typology of inventories and the requirements of universalist transcription. We might well assume that further reinterpreted insights will be achieved if they are investigated with the aid of cutting-edge phonological theory. This is what we shall do in chapter 4.

First we need to see what this phonological theory is. In chapter 3 I argue that there have been some substantial achievements in recent phonological theory and that our understanding of how individual phonological systems function and of what is generally possible in phonology is now highly developed. In the second half of chapter 3 (section 3.2) I discuss some insightful analyses proposed by others and certain novel proposals that will have a substantial impact on some of the data presented in this chapter. We will thus momentarily leave the world of diachrony for the perspective of general universalist phonology of chapter 3 before we return in, chapter 4, to the specific data sets from this chapter. This data will by no means be ignored in the coming phonological discussion, however. In fact it will drive and inform some of the analyses proposed there.

3 Theoretical phonology and lenition theory

The data that was discussed in chapter 2 is firmly established in the world of historical phonology. Although the precise details of some of the processes remain a little unclear, we can be sure that some kind of phonological processes occurred along the lines discussed and that it is reasonable to we can group them under the four names that they were given. Under the preliminary definition that was provided in section 1.4, they can all be described as featuring some kind of lenition. As we have already seen in section 2.3, the data is not as straightforward as we might hope and it clearly is the case that the precise interpretation that we give to it depends on the kind of reasoning that we use and the extent to which phonological ideas are brought to bear; even the symbols that are used to describe the processes can affect which aspects of them we see as noteworthy.

This will only get worse in the current chapter. By the end of the thesis, however, I believe that the thorough investigation that we will have subjected the data to will help us to understand them properly, in the context of a universally applicable phonology. Before this can happen, though, it will be vital to consider certain phonological issues in detail so that we have a firm footing for any claims that are made about what is a possible phonological process (and hence a possible phonological change) and so that we can come to understand what really goes on in lenition processes (and hence what went on in the Germanic lenition processes just described).

If we are to come to a full understanding of these particular processes from the history of Germanic languages, then we need to know not only what happened phonologically in each of the processes, but also we need to consider them against a general background of 'lenition theory' (as Hickey 1984 describes it). If these processes are 'lenitions' then what is 'lenition'? And how do lenition processes behave in general? For example, if some of the exceptions that the Germanic processes show are in fact *normally* exceptions in lenitions, then we might well consider them to be nothing exceptional in the Germanic case. The fact that they can be seen in the Germanic data from chapter 2 would actually be almost expected. We can only hope to discover the truth here if we dissect the notion of lenition, at least in part by considering other examples of attested lenition processes. This chapter provides this kind of argumentation and in itself thus represents a major part of the thesis, providing the general understanding of the phenomena which can later be applied to the specific Germanic data.

There are thus two halves to this chapter. The first half, section 3.1, is an investigation of recent advances in the understanding of phonological structure and phonological processes. In particular, considerable space is devoted to a theoretical discussion of subsegmental structure

because this will be of key importance to any understanding of lenition. Some of the main ideas which are adopted there have been developed in phonological traditions which see subsegmental units as privative (eg, Dependency Phonology, Government Phonology) but many are quite widely shared in phonological theory. Particular emphasis is placed on the units needed to account for the laryngeal specifications of segments. The connection with lenition processes here is obvious given that these account for the difference between series of segments such as /p, t, k/ compared with /b, d, g/. This 'phonological' section begins with a short justification of the general approach to phonology which is taken in the chapter (and in this thesis as a whole), although this needs to be tempered with a diachronic perspective, as will become clear.

The second half-chapter, section 3.2, is a detailed investigation of theories of and generalisations concerning obstruent lenition. Many of the simplifying assumptions made in section 1.4 are reconsidered or rejected and a range of previous proposals which have been made in the field of 'lenition theory' are examined; finally a novel interpretation of the concept is proposed. To give the claims that I make a wider typological base, I introduce further data in the form of a relatively brief discussion of lenition phenomena from a range of languages. These are also used, along with everything else, to propose a synthesis of lenition phenomena which takes into account both attested historical data and theoretical phonological understanding, particularly in connection with a clear comprehension of possible laryngeal contrasts.

A key part of this broad picture of lenition is played by a notion of 'lenition inhibition' which is developed in section 3.2.3. This accounts for some of the major generalisations that can be made about the phonological environments in which lenition does or does not occur.

3.1 Theoretical phonology

The classical philological approach which I described in section 1.2.1 was the main filter for the initial presentation of data in chapter 2. This 'Handbook' style is not uninformed phonologically, and when the standard Germanic Handbooks were first written, in the late nineteenth century, they were the linguistic state of the art. The sections on phonetics and phonology¹ were informed by the latest understanding of the concepts involved.² However,

¹ The distinction of approach between the two disciplines had not yet emerged, of course. It was only in the course of the twentieth century that the particularly phonological perspective developed (see, for example, Fischer-Jørgensen 1975, Anderson 1985, Robins 1990, Goldsmith & Laks 2000).

² The first volume in the series *Bibliothek indogermanischer Grammatiken* ['Library of Indo-European Grammars'] which issued a range of Handbook-style historical Grammars was Sievers's *Grundzüge der*

understanding has developed over the centuries and, although the Handbooks have frequently been reissued in new editions, the basic treatment of historical phonology remains the same. It would be untrue to say that all existing work on the processes treats the data in a Handbook-style manner, however. There have been many analyses of at least some of the data from chapter 2 throughout the twentieth century, many of which can be identified as examples of theoretical historical phonology, as defined in section 1.2. Some of the data is practically unknown to historical linguists, however. Also, while certain previous phonological analyses of parts of the data are insightful, it seems to me that they have nonetheless missed some important points.

Certain recent developments in our understanding of phonology can contribute substantially to the interpretation of the processes described in chapter 2, as I hope to show in chapter 4. These and other, less recent, but still well established tools of phonological analysis can also help us to realise how lenition processes in general can be understood. In this section I set out what they are.

It is important for a coherent theory of phonology to be based on a sound metatheoretical foundation and I devote some space here (section 3.1.1) to providing that basis, which helps to shape my understanding of what a model of phonology should look like. After that, a short section (3.1.2) justifies the use of the notion ‘segment’ in phonology, and then the next two sections deal with the two key areas of phonological representation: melody and prosody. The distinction between the two is pronounced and is consistently met in the literature, where it is also referred to as a distinction between ‘segmental’ and ‘suprasegmental’ phonology. Harris et al. (1999) is a recent discussion of some issues which help show the distinction between the two and McMahon (2001) presents evolutionary evidence which may help to place the distinction on an unshakeable foundation. Both aspects of phonology are discussed here because they will both help inform later discussion, but a clear emphasis is placed on melody as this will be of more importance later. The phonological discussion here thus deals with aspects of representation. A little attention is also devoted to theories of the nature and number of phonological levels. No real attention is devoted to a theory of derivation, however, because this is conceptually separable from the other aspects of phonology and it will not be crucially important in what follows.

Phonetik [“Basic Phonetics”] which became one of the classic standard works on phonetics in German. It had as its subtitle *zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen* [‘an introduction for the study of the phonology of Indo-European languages’]. The work was highly perceptive phonetically and very influential (see Kohler 1981). Lehmann (1993, 6) credits Sievers as being “probably the most brilliant” of the neogrammarians.

Section 3.1.3 presents a set of sub-segmental units which is in line with much current thinking and will enable an insightful treatment of obstruent lenition. After presenting a set of units which account for the place and manner of obstruents, it features a sub-section which deals exclusively, and in some detail, with the laryngeal specifications of segments. Section 3.1.4 deals with certain aspects of syllabic structure which will be relevant later. Section 3.1.5 puts all the information in the preceding sections together to illustrate the occurrence of elements in segments. These sections cannot hope to deal with absolutely all the issues that are raised in phonology, nor properly with the details of the interaction between the two key areas of prosody and melody, but a small set of coherent proposals are defended. Where points are simply borrowed from others, the justifications that were proposed by their originators are often simply assumed, as is common and necessary given the constraints of space. Section 3.1.6 concludes.

3.1.1 A minimalist basis for phonology

The model of phonology which I adopt here is naturally informed by work in established phonological frameworks, and it is clear that many of the ideas that I work with are borrowed from others. However, I do not simply take a phonological framework ‘off the peg’. When I lay out the tenets of the phonological model with which I work, I generally seek, in passing, to justify them. In section 1.2.2.2, I stated that I attempt to work with a ‘simple’ model of phonology. This is for good reason, as we will see in this section. Harris (1994b) seems to express a similar sentiment and describes his approach as ‘Minimalist Phonology’ which surely sounds better than ‘Simple Phonology’. While there are substantial differences between the model of phonology proposed by Harris (see especially Harris 1990, 1994, 1997) and the approach that I adopt here, there are also many similarities.³ These include the fact that the number of phonological levels and types of processes that are recognised are minimal, so the name ‘minimalist’ fits. There are connections here with the notion of ‘naturalness’ which was discussed in section 1.2.2.2, but in principle, the two are distinct: naturalness has to do with process-innovation, minimalism with synchronic phonological analysis, both conspire to keep phonological analysis non-abstract, however. The notion of naturalness means that processes

³ Where these similarities exist, it largely derives from the simple fact that I have taken over Harris’s argumentation and analyses.

are non-abstract at their innovation, the notion of minimalism means that they never become very abstract in synchrony.⁴

The minimalist set of assumptions that I assume here (which I describe directly) are, I believe, supported by recent work on the philosophy and phylogeny of phonology (although it must be recognised that my conclusions are not quite those of the original thinkers). One strand of thought, typically referred to as the ‘Representational Hypothesis’ (see Burton-Roberts 1994, Burton-Roberts & Carr 1999, Burton-Roberts 2000, Carr 2000) holds that the relationship of phonology to a strictly internal Chomskyan conception of language cannot be such that phonology is part of the same type of linguistic system as syntax, for example. If phonology connects with the mind external world of phonetics, then the relationship between the underlying and surface level of phonology cannot, strictly speaking, be one of realisation,⁵ ie, the entities at the underlying level cannot be linguistic in the same mind-internal way that syntactic entities are. One of the ways that syntactic entities are clearly linguistic on this picture is that they are provided by a genetically-inherited Universal Grammar. The implication is that phonology cannot be provided for by UG.

There are several other arguments for this position, but a further, relatively simple one is that the Poverty of the Stimulus argument, which is one of the best arguments for the idea that a UG plays a role in the acquisition of syntax (see, for example, the papers in Hornstein & Lightfoot 1981 and also Lightfoot 1991, 1999), does not hold for phonology (see Carr 2000), in that we are presented with evidence for the complete phonological system of a language all the time, because the phonological system is more compact than is the case for syntax.

There are substantial possible implications of this, for other branches of linguistics, as well as phonology, but we cannot contend with them all here (see Burton-Roberts 2000 for preliminary discussion). What I take from this line of argumentation is that, whatever phonology there is, it is unlikely that the sometimes baroque theoretical machinery proposed in certain phonological models (extrinsic rule ordering and ‘free rides’ in Chomsky & Halle 1968, for example, or OT sympathy theory in McCarthy 2000) can reflect the psychological reality. This is because, realistically, these could only be part of phonology if they were provided for by a part of a genetic Universal Grammar. This ties in with several ideas that

⁴ Typically, opaque phonological generalisations will be lexicalised, although if they are easily recognisable in acquisition, this need not necessarily be the case.

⁵ Instead of a literal relationship of realisation, it is proposed that phonology *represents* an internal linguistic system. I retain the traditional term, however.

have been identified (for example in Carr 1993) as general motors in the development of phonological theory since Chomsky & Halle (1968): the reduction of abstractness in phonological representations and the removal of the extrinsic ordering of phonological processes.

McMahon (2001) has presented a picture which is partly similar to that of Burton-Roberts and Carr, albeit reached by different means. She argues that there is a range of evidence which would lead us to believe that human beings have made use of prosodic phenomena for a much longer period in their evolutionary history than they have segmental melodic phenomena (aspects of prosody have clear affinities with non-human call systems, for example). This could be taken as evidence for the assumption that there has been evolutionary time for aspects of prosody to become hardwired into humans, but not for aspects of melody.

This type of argumentation does not mean that there can be no phonological universals. Nor does it mean that we should not expect to find the same type of phonological patterns in the languages of the world. It is perfectly possible that the interaction of our common human physiology with acoustics and with other factors which are common to the human experience of speech can produce hard universals which are exclusively phonological in nature. It is also perfectly possible that general aspects of mental organisation, when applied to phonology, produce certain unique effects that we might expect to be found in the sound systems of languages.

Whatever the conclusions of others from these points, mine is not that there is no phonology at all (the evidence seems to be entirely compatible with the idea that phonological processes exist), but it does seem clear to me that many processes will have a short existence in a language and their effects will be reanalysed into the lexical representations that learners construct, in line with the points made in chapter 1, much more quickly than has been assumed in more 'standard' generative historical analyses (for which, see Bynon 1977, Kiparsky 1982, McMahon 1994).

Hence I feel justified in using a minimalist phonology. To merit the name *phonology*, it will need to consider the patterning of processes and the nature and effect of a phonological system; it is also important to consider the degree to which proposed phonological entities can be seen to be active in the phonology of a language, as this is the best evidence for their existence. The phonology that I adopt here is 'post-generative' in the sense of Goldsmith (1999, 2) in that I see developments in the phonological tradition that have sprung up following, and in reaction to Standard Generative Phonology, as highly influential. I adopt an 'autosegmental', nonlinear approach.

As we saw briefly in section 1.2.2, I recognise two ‘levels’ in phonology. It is common to assume at least two levels, although many models assume more,⁶ even in approaches which overtly deny it. To my mind, the very recognition of phonological processes in a model means that there are at least two levels, even if these do not have particular formal status and particular characteristics of their own which may serve to differentiate them. It is clear in standard generative work that the levels of underlying representation and surface representation are recognisable. This also seems to me to be the case for OT’s inputs and outputs. While there is a clear conceptual difference between these two views, much of the conceptual import is the same. In standard versions of OT, the notion of the ‘richness of the base’ should affect this (‘richness of the base’ is briefly investigated in the next section). The status of inputs in OT is unclear, but they do matter to the extent that the input-output faithfulness constraints are used (for example *IDENTITY* constraints - see McCarthy & Prince 1995 - which can force input and output to be exactly the same when ranked high enough). Whatever the precise interpretation of input and output, to the extent that there can be a difference between the item in slant brackets and the chosen output candidate (even if the phonological process occurs along with many others in GEN), there can be seen to be two levels to the model.

Even if a model recognises no formal distinction between levels, in that they are both thought to be of the same ontological status, so long as it recognises that phonological processes exist, which can cause an underlying segment to surface in a form different to the stored, underlying form, then the model provides for a ‘start of phonology’ level and an ‘end of phonology’ level. I take this latter approach as a minimalist minimum, and in recognising the potential for difference between different forms of a segment, I talk of ‘underlying’ and ‘surface’ form, partly in deference to terminological tradition.

Once the case for (at least) two levels is conceded then it is true that some theoretical mechanism is required to link the underlying form to the surface form in the derivation. It is well known that various mechanisms have been proposed for this, for example, rewrite-rules, parameterised principles and ranked violable constraints. To a fair extent, the details of the mechanism can remain covert here. As was briefly discussed in section 1.2.2, the distinction in modern phonology between theories of representations, of derivations and of levels (as discussed in, for example, Anderson 1985, Ewen & van der Hulst 2001) allows us to focus on

⁶ From example, the Lexical Phonology model assumes several phonological strata which can be the locus for phonological processes and which have been ascribed distinct phonological characteristics (see, for example Kiparsky 1982, Mohanan 1986, Bermúdez-Otero 1999, McMahon 2000a).

some of these three aspects of phonology, to the exclusion of others. In a similar vein to Ewen & van der Hulst (2001) and much other work (see their references for details) I focus chiefly on issues of phonological representation.

Many of the other basic phonological assumptions that I subscribe to will be presented in the next two sections, but they can be summarised briefly here: (i) prosody and melody are different, (ii) segmental melody is composed of privative 'elements', (iii) purely phonologically motivated processes can spread these elements locally (in an 'assimilation') or long-distance (in 'vowel harmony') or delete them, (iv) such processes will thus typically be explicable in terms of their phonological environment, although (v) phonetic effects, such as the pressures of articulation and acoustics, can also affect phonology, thanks to the medium of the phonetic signal and (vi) these effects can be captured in phonology to motivate phonological processes (doubtless at least in part through acquisition, a diachronic effect); in conclusion, (vii) any phonological process can be expected to be minimal.

There are clear resonances between some of these principles and those of many well-established models of phonology, such as Government Phonology (see, for example, Kaye, Lowenstamm & Vergnaud 1985, 1990, Harris 1990, Charette 1991, Brockhaus 1995a,b, Honeybone 1999) and Dependency Phonology (Anderson & Jones 1974, Anderson & Durand 1986, Anderson & Ewen 1987, Ewen 1995) and as will become clear, there are also connections with such models as Articulatory Phonology (Browman & Goldstein 1986, 1992, Azra 1994, Hind 1996). There are also clear connections with many other linked approaches, such as Autosegmental Phonology (eg, Goldsmith 1976, 1990) and models of Feature Geometry (eg, Clements 1985, Sagey 1986). The role given to articulatory and especially acoustic effects, which will only really become clear in the course of section 3.2 has connections with the work of Ohala (eg, 1974, 1992).

The discussion in section 1.2.2.2 of the role of naturalness in the innovation of phonological processes is clearly highly relevant here. If the conjecture of, for example, Anderson (1981) and Hale & Reiss (2000) is correct, and historical process-innovation is a true locus of naturalness, then we might expect at least some common historical innovations to be easily modelable.

It is encouraging to note that other kinds of phonological thinking are converging on the position of a relatively minimalist phonology, even if from very different starting points. We have already noted that work in Government Phonology allows only a very restricted set of processes and rejects one type of abstract phonological analyses in a principled way (see, for example, Kaye 1995). It is also true that certain writers in Optimality Theory have rejected the

complications of analysis which are brought in by theorists in order to deal with abstract analyses⁷ (for example, Hammond 1999 discusses this idea, see Honeybone 2000a).

There is one remaining facet of the model of phonology that I adopt here which may require some justification. This is the extensive use of the notion of the phonological segment. Segments were employed without comment in the discussion of the data in chapter 2, in line with the kind of presentation found in the Handbooks. As intimated above, I make explicit use of subsegmental structure in coming chapters, but I also continue to refer to the segment. This might be contentious from the standpoint of certain contemporary models of phonology, so I turn to this question briefly in the next section, before proceeding to the modelling of melody and prosody

3.1.2 Phonological segmenthood

In preceding discussion I have made much use of the concept of the phonological segment and I shall continue to do so throughout this thesis. In the phonology of recent years, two separate developments have called the notion of the segment into question, however. These proceed from very different perspectives and the argumentation is different in each case, but I believe them both to be mistaken in this conclusion, although certain parts of the argumentation that has led some to this belief is certainly correct. The practise of many phonologists seems to provide corroboration for my approach, in that the term *segment* still plays a major role in phonological discourse and it is not clear that many phonologists have actually abandoned the concept.

The two developments are largely separate from each other, although they are in principle compatible. They are Autosegmental Phonology and Optimality Theory. The challenge to the notion of the segment from Autosegmental Phonology (for the development of framework see Goldsmith 1976, 1990) comes from the formalism and reasoning which has been developed within that general research programme (which has been widely accepted in phonology and finds expression in a large number of frameworks, often being referred to as ‘nonlinear’ phonology, see, for example van der Hulst & Smith 1982 and Anderson 1985, Goldsmith 1999). The major insight behind this kind of analysis is that subsegmental units (normally referred to as ‘features’) can be shared by segments. The idea was first proposed in generative phonology by Goldsmith (1976) to account for tonal phenomena which can be seen quite

⁷ See McMahon (2000b) for a discussion of many of the types of theoretical machinery that been proposed in the broad church of OT literature.

clearly to spread over more than one segment, but it was quickly extended to account for subsegmental processes, such as assimilation, where a feature can simply be attached from one segment to a neighbouring segment, meaning that the two share the feature in question.

In some models of phonology (including that used here) all subsegmental units are considered to have the status of autosegments. This means that they can attach across the boundaries of what is traditionally considered to be a segment. Goldsmith (1990) writes “the individual gestural components of articulation - the features of modern phonology - each have quite separate lives of their own, and an adequate theory of phonology will be one that recognizes this, and provides a way to understand the linkages between individual gestures of the tongue, lips, and so forth, and larger units of organization, such as the syllable”. Much of the reasoning behind this is highly compelling, but, if features have a life of their own, where does this leave the segment, which we might want to define as a particular combination of features? If features are spread all around a word and can attach directly to syllable nodes, the notion of the segment is in danger of disappearing.

The idea that the most basic units of phonology are smaller than the segment and can be manipulated separately will be of vital importance in this thesis, but it does not necessarily mean that there is no segmental level of organisation as well. Without segments, we cannot have segmental inventories, nor phonological systems as they are normally considered. The segment does seem to play a clear role in phonological processes, for example in cases of metathesis where whole segments are moved by the phonology of a language. Without the notion of a unitary segment, it also becomes difficult to understand phonological changes which spread to all occurrences of a segment in a language, having started by affecting occurrences of the segment in only one particular phonological environment. Some notion of analogy seems to be at play here, but it crucially rests on the psychological unity of a segment in a language. It strikes me that this is a vital aspect of many models of phonology (even if it is often not overtly recognised) and it will be important in later discussion.

There is also substantial evidence from psycholinguistics that segments are involved in both speaking and hearing (see Cutler 1992, Ohala 1992) and it would be odd if the segmental entities needed there do not map onto real units in phonology. It seems that the notion of segment is a vital one and I continue to use the concept below. In section 3.1.5, we will note a simple phonological definition of a segment which allows us to recognise autosegmental behaviour in subsegmental units and the segment as a unit as well.

In the Optimality Theoretic literature, a further problem arises for the notion that underlying segments exist. This is typically known as the ‘Richness of the Base’ (Prince &

Smolensky 1993) and it derives from the fact that in standard OT there can be no constraints on inputs, which, as we saw in the last section, leaves the input in an uncertain position theoretically. If there is no clear notion of which kinds of segments can occur in inputs, then the notion of underlying segment, such as can be used to construct inventories, again appears on shaky ground (as, eg, Kirchner 1997 notes, contrastiveness is a facet of constraint ranking in strict OT).

As McMahon (2000, 36) reports, however, it is not the case that all work in OT adopts a 'Richness of the Base' approach, where inputs are unimportant. Archangeli & Suzuki (1997) adopt an input markedness constraint, for example, which can be taken as an indication that something is amiss with the utter lack of attention that the input receives in OT. Of course, the same arguments as above still hold here for the psychological existence of the segment, against the standard OT position. I make substantial use of the notion of the segment in what follows, and I feel this to be fully justified.

3.1.3 Melody: elements and privativity

This section focuses on the crucial aspects of phonological representation which will help us to clarify and classify the diachronic phonological data discussed in this thesis. I first explain the underpinnings of the approach adopted, which is in general compatible with and hence dictated by the 'minimalist' assumptions which were outlined in section 3.1.1. I then present a model of subsegmental representation, partly in this section and partly in a specialised subsection (3.1.3.1). A summary section (3.1.5), which follows a brief discussion of prosodic concerns, shows how these elements feature in segments.

The discussion here will inform and be informed by the investigation of the Germanic processes from chapter 2 in particular and of lenition in general. Particular questions concerning lenition are addressed (in section 3.2) after the current, more generalist discussion and so will not feature heavily here. The kind of argumentation used in this section will, however, on occasion result from issues arising from the study of lenition. Because the focus of the thesis is purely on obstruents, I only dwell on points which are relevant to the description of this kind of segment. I briefly deal with some issues that are relevant to the description of vowels and other types of segments, but not in any detail.

Despite the space devoted to the justification of the retention of the notion 'segment' in phonological theory in section 3.1.2, the idea that segments are composed of smaller units will play a vital role in coming discussion. Goldsmith & Laks (2000) show that the

'discovery' of such units has been seen as perhaps the greatest insight in twentieth century phonology, compared by Jakobson (1976) to the discovery of the ultimate constituents of matter in physics.⁸ The precise nature and number of these subsegmental units has probably been the subject of much greater debate than has surrounded the discussion of atomic and subatomic particles, however. The name given to the units also varies: *features* is by far the most common, but they have also been described in differing phonological models as *oppositions*, *components*, *particles*, *gestures* and *elements* and probably other things, too. I use the last of these names, for reasons which I give directly, but the original insight is maintained in practically all phonological models and I argue below that the differences between such models are not as great as is sometimes supposed.

Phonological features for obstruents are often split up into three main kinds: (i) those which account for the place of articulation, also known as 'point of articulation' and 'location', (ii) those which account for manner of articulation, also known as 'stricture', and (iii) those which deal with laryngeal articulation, also known as 'voicing' and 'phonation' or 'initiation'.

In this section I first engage in a general discussion of points which inform and justify the type of features that I assume, and then proceed to a description of those features which I assume for place and manner. I devote the following specialised subsection to laryngeal specifications because they will play an extremely important role in the discussion of lenition, especially in the Germanic data.⁹ Certain aspects of what is traditionally referred to as 'manner' will be shown to be a facet of the interaction between the melodic units proposed here and aspects of prosody dealt with in section 3.1.4.

Features are elements

In section 3.1.1, I argued that a model of phonology should be minimalist. Such a minimalist model of phonology requires a minimalist theory of features. This principle could be reflected by the model in several ways: (i) by using only a small number of features (although not so small that the features would need to be manipulated with extensive theoretical machinery),

⁸ Jakobson's own role in the initial formalisation of the insight of subsegmental structure (eg, Jakobson, Fant & Halle 1952), along with that of his Prague Circle colleagues such as Trubetzkoy (eg, 1939), is well recognised.

⁹ It is naturally true that the system of features that a theorist assumes could have an impact on their understanding of segmental processes such as lenition, to the extent that there are real differences in feature systems. When we arrive at the discussion of lenition in this thesis, the feature system that I assume will play a role in my evaluation of competing accounts for lenition processes. I am thus pre-empting some of the discussion of theories of lenition there by arguing for the model of subsegmental phonology in this section before I present existing ideas of lenition theory, but either the chicken or the egg must come first.

(ii) if the features essentially function in the same way as each other (rather than having some features with one set of characteristics and others with a different set), (iii) we would need to reject a model which requires abstract ordered phonological processes simply so that the features can be interpreted, and (iv) the principle would also be met if features can be shown to have some correlation with phonetic units.

Luckily, there are existing phonological models which go a long way in fulfilling these criteria. The model that I adopt here is inspired by several phonological traditions, some of which are linked, and all of which are compatible with these ideas to a greater or lesser extent. I do not take over any model completely, but the proposals that I make are inspired particularly by work in Dependency Phonology (Anderson & Jones 1974, Anderson & Durand 1986, Anderson & Ewen 1987, Ewen 1995) and Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990, Harris 1990, Brockhaus 1995a,b), with resonances in the allied work of Rennison (1986, 1990), Particle Phonology (Schane 1984, Broadbent 1999), Articulatory Phonology (Browman & Goldstein 1986, 1992, Azra 1994, Hind 1996) and even Autosegmental Phonology (eg, Goldsmith 1976, 1990) and models of Feature Geometry (eg, Clements 1985, Sagey 1986).

The approach that I take probably finds its roots in ideas that were first formulated in Dependency Phonology, where the subsegmental units are called 'components', but, in common with others, I refer to it as 'element theory', following the terminology of Government Phonology which has now also been adopted in some Dependency literature (eg, van der Hulst 1995). In the coming discussion, I occasionally use the word *element* where an original author may have used a different term. It is not the case that all researchers accept exactly the same ideas, however. To be explicit, the principles which I take from this approach are as follows: (i) subsegmental units are elements, (ii) there is a relatively small number of elements, (iii) elements are exclusively privative, (iv) elements behave in an autosegmental manner to the extent that they can spread to cause assimilations and longer-distance harmonies, (v) elements have the capacity to be interpreted without support from other machinery, ie, some segments consist of only one element, (vi) elements in one segment can contract simple relationships with each other. This set of assumptions, the segmental representations and types of phonological processes that it allows are quite minimalist as we will see, yet a range of insightful phonological analyses have been proposed in the broad tradition from which they spring.

Some of the implications of these principles are that segments can be built up from several elements (either statically or by phonological process) or could lose an element if they have

more than one element in their make-up; the loss of the only element in a segment would result in the loss of the segment. Thanks to these ideas, it is possible that phonological processes can be used as evidence for the nature of elements because they can reveal the structure of segments. To an extent, elements can be seen to group together the features of approaches that work with a relatively large number of (often binary-valued) features, although this is not typically thought to be a psychologically real description.¹⁰ In my description of subsegmental units, I call them *elements*, after work in Government Phonology, and when I write them, I enclose them in vertical slashes (thus: |element|), following a convention of Dependency Phonology.

The initial insight for the approach came from the analysis of vowel systems, where it is noted that the most basic phonological system is one with three ‘corner vowels’ /i, a, u/. This seems to be the most unmarked system in language inventories in that they are the most common vowels in languages and they represent the extremes of the vowel space. It is proposed that these vowels consist of only one element: |I|, |A|, |U|, respectively. These elements are directly phonetically interpretable and do not ‘alter’ in any way during a derivation. Other vowels are made up of more than one element (so, for example, the vowel /e/ consists of |I| and |A|, the vowel /o/ consists of |U| and |A|, and the vowel /y/ consists of |I| and |U|). This contrasts quite distinctly with a ‘standard’ distinctive-feature position, such as that found in Chomsky and Halle (1968), where the segment [u], for example, crucially has the specification [+high] and [+back], but in order for these features to be phonetically interpretable, it requires a host of other specifications, some predictable (or ‘redundant’), some not.

The elements are often glossed to describe what they add to a segment, and this gives them a correlation with phonetic properties. For example, Anderson & Ewen (1987) gloss |I| as ‘frontness’, |A| as ‘lowness’ and |U| as ‘roundness’.¹¹ It is generally recognised that other

¹⁰ In the initial exposition of Government Phonology (Kaye, Lowenstamm & Vergnaud 1985); this approach was in fact explicitly adopted and certain elements were described as being composed of a set of feature values, but this is not normally encountered now. I also do not consider the concept of ‘charm’ which has, at times been associated with elements (see Honeybone 1999 for discussion).

¹¹ There is a debate in the literature as to whether elements should be interpreted as units which are mapped primarily onto articulatory gestures or acoustic properties. Elements are mostly given articulatory definitions, as with Anderson & Ewen (1987), and many others such as Smith (1988) and Yoshida (1996), but others, for example, Harris (1994), Harris & Lindsey (1995) and Brockhaus (1995) propose that the elements should be thought to map directly onto acoustic properties of the signal. In this way, Harris (1994) glosses |I| as “low first formant coupled with a spectral peak (representing the convergence of Formants 2 and 3) at the top of the sonorant frequency zone” and |A| as “a spectral peak (representing the convergence of Formants 1 and 2) located in the middle of the sonorant frequency zone” (1994, 140). Ideally, elements would map onto both articulation and acoustics, and Brockhaus (1995) in fact gives glosses for both articulatory and acoustic properties. In what

elements are needed to construct vowel systems which are more complex than the basic three vowel system (for example, Harris 1994 and those following him also use an element |@| which can be interpreted as ‘neutrality’ so that the combination of |I| and |@| gives /ɪ/).

The approach has been extended in various ways to account for consonants as well as vowels and this has been a fruitful source of debate in the literature. There have been many proposals for precisely which set of elements is needed and I adopt a quite simple set below. It is widely accepted that glides such as /j/ and /w/ consist simply of the elements |I| and |U| respectively, situated in a non-nuclear position (we return to the notion ‘non-nuclear position’ in the discussion of prosody in section 3.1.4). The representation of obstruents is more complex, however, and I present a set of elements to capture obstruent structure below. Before that, I briefly turn to a few more issues that are raised by the use of elements to capture subsegmental structure: the advantages of privativity and the lack of underspecification, and the notion of headedness.

The claim of privativity is that an element is either present in the representation of a segment or absent, in which case it can have no effect; the notion of, for example, ‘[– back]’ or rather ‘[– |U|]’ is not expressible in the theory. The notion of privativity goes back to Trubetzkoy (1939), who proposed that only some features might be thought of as privative. A good deal of phonological work in the ‘standard’ feature and ‘feature geometry’ tradition (see, for example Sagey, 1986, Lombardi 1994, Clements & Hume 1995, Steriade 1995) assumes that certain phonological features are privative. In element theory, as mentioned above, all elements are proposed to be privative (also described as ‘monovalent’, ‘unary’ or ‘simplex’). This is clearly a more minimalist assumption and it seems fair to think it is the null hypothesis. If some features are privative, why not all?

At times, the effects of the difference between the positive and negative value of a binary feature can be translated into privative features in that the positive value could be represented as |elementA| and the negative value as |elementB| but by itself this does not speak for the superiority of either approach (and it is not normally the case that such translation be done). There is a conceptual difference between using two elements such as |palatality| and |dorsality| and using two values of a binary feature such as [–back] and [+back]. Elements can often be connected with some positive effect such as movement in the tongue towards a particular goal

follows I use articulatory glosses (very overtly at times), partly because it seems clear to me that the pressures and possibilities of articulation do affect the phonology of a language. It does not seem an impossible long term goal, however, that the two types of glosses may be unified. A further note: the term ‘gloss’ used here is largely equivalent to the notion ‘hot feature’ in Kaye, Lowenstamm & Vergnaud (1985).

and this is not always the case with binary features. It seems uncertain why a feature such as [-back] should be read as an instruction to make an articulator move in some positive way (for example, the tongue towards the front of the mouth), rather than as an absence of an instruction to move towards the back of the mouth. It is conceptually problematic to propose that something positive is achieved (tongue advancement) through a command not to do something. All this is avoided through the use of privative elements.

A clear consequence of the use of this type of primitive is that there is no notion of underspecification in lexical representations (or anywhere else). This is an important idea in elemental approaches and is clearly consistent with the approach adopted here because underspecification analyses require often numerous externally ordered ‘default rules’ to fill in lexically unspecified values which remain at the end of the phonology (this is especially the case in Radical Underspecification, which strips out all possible non-contrastive information from lexical entries). As many have pointed out (eg Goldsmith 1995, Steriade 1995, Harris & Lindsey 1995, McMahon 2000a) underspecification relies on the questionable notion that lexical storage capacity is at a premium and phonological derivations can be unlimitedly complex (Steriade 1995 calls this ‘lexical minimality’). It also requires a complex mechanism (Goldsmith 1995 calls it ‘device *D*’) which needs to calculate which feature would be strictly speaking unnecessary or unmarked for each pair of segments in each possible environment. As McMahon (2000a) points out the proposed model would not be easily learnable. The idea of underspecification is diametrically opposed to the minimalist position taken here. Those who propose underspecification have not provided evidence for the assumption of lexical minimality and device *D*. There is no evidence that human beings have a particularly limited amount of mental storage space available and, as we have seen, there is evidence that speaks against the assumption of overcomplex phonological models.

The notion that elements are in principle ‘independently pronounceable’ severely restricts the prospect for underspecification and even where this is in doubt, there is little reason to replace it with underspecification. The arguments for underspecification are currently under attack from all fronts, including the prevailing current in Optimality Theory (see, for example Kager 1999).

The embracing of privativity and the lack of underspecification make the model adopted here unlike certain phonological models but very much akin to others. As we have begun to see and as will become clearer soon, the set of elements used, along with other aspects of the phonological model, are not precisely the same as any other. Nonetheless, there are substantial similarities between several analyses proposed here and counterparts from other

frameworks. As Ewen & van der Hulst (2001) note: “in recent years it has become apparent that many claims made in the various models are not in fact independent of each other, and that claims made within the framework of one approach are often restatements of those made elsewhere” (2001, xi). Many analyses made in one framework are translatable into other models. It will be quite clear how some of the analyses proposed here are compatible with Government Phonology and Dependency Phonology, but many of them are compatible with others, too. In fact, some of the key proposals and positions that I defend are inspired by work which has been carried out at least partly in very different basic frameworks but which I see as being entirely translatable into Element Theory (and the translation back again is equally possible, of course).

For frameworks that use privative features or gestures, parts of this translation will be relatively straightforward. While it is not to be taken literally, the ‘gathering of features’ into elements, as mentioned above, makes certain elements function like certain nodes in feature geometry trees. This is especially true for place features and laryngeal features, which are often viewed as privative even in approaches which allow a mixture of feature types. The autosegmental-type spreading and sharing of elements is, in any case, basically the same process as that used very widely.

There are also resonances with Articulatory Phonology, where the subsegmental primitives are articulatorily defined ‘gestures’. While the kinds of gesture manipulation allowed in the standard version of the theory (Browman & Goldstein 1989, 1992) are not the same as those countenanced here (for example, elements can be fully lost in a process, while they can only be obscured by other gestures in strict Articulatory Phonology), the primitives are not entirely unlike. The ‘gestures’ of Articulatory Phonology are privative, in that they are conceived of as a kind of positive command to articulators (which is one possible interpretation of elements) and underspecification is not countenanced because gestures cannot be inserted in a process, which is similar to the stand adopted here in that processes must originally be motivated by their phonological or phonetic environment. Finally, work by McMahon, Foulkes & Tollfree (1994) and McMahon (2000a) has sought to ‘phonologise’ Articulatory Phonology, with the effect that some of the gradient phonetic aspects of the gestures are removed from a strictly phonological level. Every model of phonology requires a type of phonetic implementation at which gradient effects are visible and it is possible that this level would be similar for standard Articulatory Phonology and a phonology with elements. Indeed, the commensurability of Articulatory Phonology with elemental approaches has not gone unremarked before (see Azra 1994). We return to certain aspect of this phonological model in

section 3.3.2.2.4, where the focus of discussion is the modelling of lenition processes in phonology. Naturally, certain other important aspects of phonology are discussed in that section as well.

There is one final crucial aspect of the elemental model that I adopt here which requires attention. This comes to light when elements combine to form segments, as we will see in section 3.1.5. When a segment consists of only one element, then there is little else to say because this does not allow for any kind of relationship to be contracted between elements. However, when elements combine, the possibility arises that the elements may interact. This is especially important when two segments which can contrast in languages contain precisely the same set of elements, which is not an unusual occurrence given that a relatively small number of elements is used, and, in fact, it is a desirable result, as we will see. When segments consist of more than one element, there are two possibilities: (i) the elements could function in an equally important way in the make up of the segment or (ii) they could be ranked in some sense in terms of the relative contribution that they make to the segment. This latter idea has been widely taken up in the literature and has intuitive appeal. I adopt it here.

The insight is characterised in Dependency Phonology by the identification of dependency relations between elements (see Ewen 1995 for an investigation of the Dependency Phonology notion of ‘dependency’ and comparison with similar ideas in other phonological models). In Government Phonology, it is normal to identify one element in a segment as the ‘head’ of the segment, with any others being dependents. This is a straightforward mechanism and, where it is necessary, I adopt this approach below. In the segmental representations to come, a head is only recognised where there is evidence that more than one segment consists of the same set of elements; it is, in fact, only here that the notion of headedness is needed. This mechanism is not ad hoc. It allows the theorist to retain a relatively small (‘mimimalist’) set of elements and to express the reality of the situation where two segments, for example /θ/ and /s/, are practically identical, but differ in some small regard. It seems reasonable to characterise this as a slight difference in the prominence of one of the segments’ attributes. The notion has proved so successful at characterising something that needs to be characterised that it has also been imported into Articulatory Phonology by Hind (1996), which further illustrates the degree of translatability between the models.

It is worth noting here that the phonological head-dependent relation which is generated in this way is not necessarily the same as that proposed in other areas of linguistics. While it is well-known that formal head-dependent relations exist in syntax, it is not clear that the

relationship is really the same. As Carr (2000) points out, in syntax a complement is obligatorily selected by its head (or forbidden by it, as in the case of transitive and intransitive verbs). This seems a rather different idea to the simple notion that one part of an item is more prominent than another. The relation of headedness adopted here for phonological segments is thus really simply a case of the increased salience of one element at the expense of the others. This could easily be seen to have a basis in general cognition.

A set of elements

The corner vowel elements discussed above ([I], [A], [U]) are standardly assumed in elemental theories and some kind of [ə] is common, too. The elements that I use here for the description of consonants are not so standard in terms of element theory, however, and this is partly because they have been formulated to flag up their articulatory identity. This will illustrate the cases where they are easily comparable with other feature systems. The most important aspects of segmental structure for the study of lenition are those aspects which are traditionally referred to as ‘manner of articulation’ and ‘laryngeal specification’ and the proposals here will be explicit for the elements and representations required to capture these. As we saw in section 1.4, ‘place of articulation’ is not relevant in lenition (there is no change in major articulator), consequently the elements proposed for place will be somewhat simplifactory.

I retain the vertical slash notation, but I do not use a classical ‘element’ symbol for the following elements. I simply name their key property, enclosed in the slashes. They are conceived of as abstract mental phonological units, and their phonetic exponence is not always precisely what their name suggests, but this notation is intended to be easily comprehensible and, on the question of place, it is intended to be neutral regarding some live issues, the discussion of which would take us too far afield from the aim of this thesis.¹² We return to these issues very briefly in section 3.2.3, but there is not the space to explore them in this thesis. The remainder of this section presents a set of elements for the place and manner of obstruents.

¹² My approach has the result of making the place elements more akin to the representations of Sagey (1986) than those of Government or especially Dependency Phonology. Sagey adopts privative ‘class features’ which have the same name as the elements adopted here. She also uses terminal features, however, which are binary and are attached to the class features. This is clearly different to the approach adopted here, which captures some of these effects by a difference of headedness and views others as the effects of unitary elements.

Elements for place

It is widely recognised that there are three major places of articulation for obstruents. This correlates with the fact that the vast majority of the world's languages have series of stops, for example, with all and/or only these three places of articulation (according to Maddieson 1984, 98.4 % have stops at these three places). Three place elements are thus a minimum, and I use the simple 'named' elements given in (3.1), along with some of the types of segments that they characterise:

(3.1)

|labiality|

- in bilabials and labiodentals

|coronality|

- in dentals and alveolars

|dorsality|

- in 'back' consonants including velars

Together with the notion of variable headedness, these account for all the segments encountered in this thesis. Examples of their use in segments are given in section 3.1.5. Further place effects encountered in other languages are naturally characterised in other ways (see Anderson & Ewen 1987, Smith 1988 and van de Weijer 1996 for some suggestions).

Elements for manner

Another of the classic parameters of segmental description is 'manner of articulation'. This is much more important for the characterisation of lenition than was place and hence the proposals here are not thought of as simplifications. One of the key discoveries of non-linear phonology has been that the traditional notion of 'manner' is not really a uniform category. Gemimates and affricates are represented through the interaction of the elements of melody with the most basic unit of prosody, as we shall see in section 3.1.4. Other aspects of manner can be accounted for using the elements given in (3.2), along with some of the segment types that they characterise:

(3.2)

|occlusion|

- in stops and affricates

|frication|

- in fricatives and affricates

|nasality|

- in nasals

The set of elements used here is most clearly influenced by those developed in Government Phonology, such as the slightly different sets in Kaye, Lowenstamm & Vergnaud (1989) and Harris (1990, 1994). The facts of lenition have played an important role in the development of ‘elemental’ theories of subsegmental structure, particularly as regards elements for manner, as we will see in section 3.3.2.8. There are crucial differences in what I assume to the models just mentioned, however and these are discussed in that section.

The element |nasality| will not feature further in this thesis but it is required in phonology to account for nasal stops and nasalised vowels. The element |occlusion| represents complete closure in the oral tract at the place specified by place elements, and |frication| represents close approximation. These elements are further explored in sections 3.1.5 and 3.2.2. Before that, we turn to the investigation of elements for laryngeal specifications in detail.

3.1.3.1 Laryngeal specifications

The last section dealt with the key general issues which underlie the assumptions that are made in this thesis as to the nature of subsegmental elements. These will naturally carry over into this section, and thus the elements adopted here will display the general properties that were ascribed to elements in that section, such as privativity and the potential (at least) to be interpreted as a segment in their own right. The last section also presented a set of elements to account for those aspects of segmental structure that are traditionally described as ‘place’ and ‘manner’.

The current section is entirely devoted to the third classic parameter of phonetic classification: ‘voicing’, which I refer to as ‘laryngeal specifications’ because most of the articulatory actions that are connected with the elements which are to be proposed involve activity in the larynx. As will become apparent, this will prove to be extremely important for the reanalysis of the historical changes that is to be proposed in chapter 4, so substantial

discussion will be devoted to the issues in this section (much more than was devoted to the consideration of elements and representations for place and manner).

One of the key claims made here is that most conventional analyses of the underlying laryngeal specifications of segments in languages such as English, German and their historical predecessors have been mistaken. This is proposed in connection with a set of unconnected yet compatible recent proposals which have considered the features that are needed to account for the synchronic laryngeal phonology of contemporary languages (and as we shall see, these proposals have a substantial tradition to back them up, in fact). This will involve a general reinterpretation of how typical systems of laryngeal contrast function. A key result of the discussion in this thesis will be the application of this position (which I refer to as ‘laryngeal realism’ because I believe it makes clear the true nature of possible laryngeal contrasts) to the historical Germanic lenition data that were presented in chapter 2 and to our understanding of lenition in general. The current section is thus key to the thesis.

Because it takes on an important role in the thesis, I devote some quite substantial discussion to the topic of laryngeal specifications. The initial discussion (in sections 3.1.3.1.1 to 3.1.3.1.3) deals with the general phonological possibilities, focusing on the patterns of contrast which can be observed in contemporary synchronic languages. This will lead to a detailed application of the uniformitarian principle to an area of historical phonology which has, I think, been sorely misunderstood. Because the implications of some of the positions that are defended here are quite substantial, it will be worth going back to basics and investigating what is possible in laryngeal phonology, to avoid a focus on just one genetically related group. This is the first task of this section.

As a part of this, previous ‘standard’ approaches to laryngeal specifications are examined, after which, in sections 3.1.3.1.4 and 3.1.3.1.5, the reasoning which has led various researchers to replace the conventional wisdom in this regard with the position which I adopt here (in 3.1.3.1.6), is discussed; we will see that the proposals are extremely translatable, even though they were formulated in quite different theoretical frameworks. I then present further novel evidence for this position and begin to draw out the implications of the position for historical phonology and to investigate how diachronic evidence weighs on the issue (3.1.3.1.7 to 3.1.3.1.9).

One as yet practically unexplored issue is the question of laryngeal specifications in fricatives. As Vaux (1998) notes, practically all of the work on the topic has dealt with stops. Following Vaux’s lead, I devote some discussion here (in section 3.1.3.1.10) to the laryngeal

specifications for fricatives. This will have obvious implications for our understanding of the EIFV from chapter 2.

3.1.3.1.1 The starting point

One of the earliest forms of segmental transcription for the Germanic languages, as for most Indo-European languages, was the use of the Roman alphabet in writing.¹³ The Roman alphabet had naturally been designed to fit the phonology of Latin and when medieval monks and others came to write the Germanic languages, for example, they succeeded in introducing a few new letters to make the alphabet fit their phonology better (such as ⟨þ⟩ and ⟨æ⟩, for example) but mostly they had to take over what they inherited. In the case of stops, there was no real problem because Latin had two series of stops and all Germanic languages had two series of stops given that both Germanic and Italic had lost one of the IE series through a spirantisation,¹⁴ so the opposition ⟨p, t, k⟩ : ⟨b, d, g⟩ worked for the phonology of all these languages.

The same Roman alphabet has formed the basis of avowedly phonetic systems of transcription such as the IPA, and it is still true that the basic symbols for stops are the Roman letters, augmented by small capitals and some newly created symbols for non-Latin segments. There are two main series¹⁵ of symbols which are described as ‘voiced’ and ‘voiceless’. Outside of European languages, there are various other quite commonly described types of laryngeal specification, however. Indeed, we have already encountered some in the ejectives that Gamkrelidze & Ivanov (1973) and Hopper (1973) proposed in the Glottalic Theory, discussed in section 2.3.3. The ‘voiced aspirates’ (*MA*s) of traditional IE are also naturally interpreted as having a different type of laryngeal specification to the standard ‘voiced’ : ‘voiceless’, as are the ‘voiceless aspirate’ *TAs*, by those who still cling to them.

In the Present-Day descendants of the Germanic languages described in chapter 2, the phonological distinction between the two is typically characterised very straightforwardly: the segments /b, d, g/ are [+voice] or ‘voiced’ and the segments /p, t, k/ are [–voice] or

¹³ This is to ignore the use of runes, of course. The runic alphabet relied on the same principle as the Roman alphabet, however, and the vast majority of early writing that we have for most Germanic languages is in an alphabet which was derived from that use for Latin. It also ignores the case of Gothic, for which Wulfila devised his own alphabet (see Braune & Ebbinghaus 1981 for example), however, the general point holds.

¹⁴ A different series spirantised in each language, however; the Germanic spirantisation was GCS₁, of course.

¹⁵ It is true that voiced implosives now have dedicated unitary symbols, too: /ɓ, ɗ, ɟ, ɠ, ʄ/, but other types of stop, including the third and fourth most common types, aspirated voiceless and ejectives, (Maddieson 1984) do not.

‘voiceless’ (see, for example, Booij 1995 for Dutch, Wiese 1996 for German, Hammond 1999 for English). Equally, Present-Day descendants of (Vulgar) Latin typically have the distinction between stop series characterised in exactly the same way (see Macpherson 1975 for Spanish, for example).

3.1.3.1.2 What is possible in laryngeal phonology?

Ladefoged (1973) describes 11 different types of stop which can be distinguished by ‘phonation type’, gathered from various different languages. The notion of phonation is clearly linked to the set of laryngeal elements in that, on our model, it is they that can work alone or combine together to produce what Ladefoged labels ‘phonation’. Ladefoged’s (1973) table is reproduced here (from Lombardi 1991, 1995) as (3.3). The key question in connection with a table like this, as with much phonetic work on the issue, is to what extent do the phonetic categories reflect phonological ones (and Ladefoged is well aware of the question). While this issue is obviously of crucial importance in any phonological analysis, it seems to have been a particularly live issue for laryngeal specifications.

(3.3)

		Hausa	Korean	Hindi	Sindhi	Igbo	Uduk	Siswati	Beja
a	laryngealized (creaky voice)	x							x
	voiced implosive				x		x	x	
b	voiced	x		x	x	x	x		x
c	voiceless lenis		x						
	voiceless	x		x	x	x	x	x	x
d	murmured (breathy voice)			x	x	x		x	x
e	aspirated		x	x	x	x	x	x	x
f	voiceless fortis		x						
	voiceless ejective						x	x	
	glottal stop plus stop								x
	glottalic ingressive					x			

The table lists 11 possible stop types, and an ‘x’ indicates that the series has been described for the language concerned. This could be problematic for an element-based approach as it raises the spectre that up to 10 individual elements might be required, but this is by no means necessary. Ladefoged (1973) had already phonologised his table to indicate which kind of segments of the 11 listed ever contrast in languages. The lettered groupings of segments indicate this: those in one group never contrast. Lombardi (1991) abstracts away from this to recognise the six possible contrasting laryngeal specifications in stops and replaces Ladefoged’s features with a smaller set. The six types of stop are given in (3.4):

(3.4)

- a. voiced glottalized (usually implosive)
- b. voiced
- c. voiceless
- d. voiced aspirate (“murmured”)
- e. voiceless aspirated
- f. voiceless glottalized (usually ejective)

From this we can recognise ‘f’ as a segment type assumed in the Glottalic Theory and ‘d’ as the description for the *MAs* in the traditional reconstruction, and quite possibly in many modern reconstructions, too, as was discussed in section 2.3. Hopper (1973) explicitly describes the stop as ‘mumured’ and ‘voiced aspirate’ is arguable still Gamkrelidze & Ivanov’s (1973) reconstruction, given that it is proposed to have had contextual aspiration. Segments ‘b’ and ‘c’ are common descriptions, as noted above.

3.1.3.1.3 What elements are needed for laryngeal specifications? Take 1

Several proposals have been made in ‘feature’ theory for the characterisation of laryngeal specifications. Halle & Stevens (1971) is widely cited as an influential proposal (for example by Keating 1984, Kenstowicz 1994). Halle & Stevens propose that segmental laryngeal specifications be captured phonologically using four features, each of which is binary. Their set of features is given in (3.5):

(3.5) Halle & Stevens (1971) laryngeal features

- [± spread vocal cords]
- [± constricted vocal cords]
- [± stiff vocal cords]
- [± slack vocal cords]

The Halle & Stevens (1971) system has no difficulty in characterising the laryngeal segment types of all known languages. In fact, this is the problem. It can characterise such a wide range of possible segments¹⁶ that it predicts a number of segment types should exist that do not in fact, and also several of the segment types given above can be described in more than one way using this set of features. Keating (1984) is surely right when she writes that, in Halle & Stevens’ system “features distinguish voiced laryngealized stops from true implosives, whereas languages never do [...and they...] needlessly distinguish voiceless unaspirated from voiceless lax stops. The point, then, is that H&S (and *SPE*) don’t simply have the wrong features in these instances; they will ALWAYS have TOO MANY features because they want to describe exactly how individual sounds are articulated. While we want the phonological features to have some phonetic basis, we also want to distinguish possible contrasts from possible differences.” (1984, 289 original emphasis). Keating’s argumentation

¹⁶ The features could in principle combine in a large number of different ways, although certain combination of features are thought to be impossible on physiological grounds (see Kenstowicz 1994).

is often highly compelling¹⁷ and Halle & Stevens (1971) system is not now widely used in its entirety.

As far as we are concerned here, there is a further major problem with Halle & Stevens' proposal: it uses binary features. If we accept binary features into phonology along with privative elements then the model loses considerably in minimality. It has to be said that Halle & Stevens' system is not easily translatable into an elemental viewpoint. A relatively standard development of their system is to collapse [± stiff vocal cords] and [± slack vocal cords] into [± voice] but this only makes it a little more appealing from the element perspective. We return to Keating's (1984) own proposal briefly below, but leave it momentarily because she only deals with 'European-type' languages and does not address any issues relevant to ejectives and other less common stops.

One proposal that succeeds in countering the chief problems identified with those that we have considered above is that of Lombardi (1991). Lombardi surveys a wide range of data and proposes to account for laryngeal phonology using only three privative features. This is clearly attractive from a minimalist perspective. Her three features are shown in the top line of (3.6), which also revisits the segments from (3.4) which she and Ladefoged identified as being possibly contrastive in languages. The table shows how Lombardi links the individual features to their occurrence in segments; the plus signs are simply an indication of the presence of the feature in the make-up of a segment:

(3.6)	[voice]	[glottalization]	[aspiration]
voiceless			
voiced	+		
voiceless aspirated			+
voiced aspirated	+		+
voiceless glottalised		+	
voiced glottalised	+	+	

This seems to capture the facts in an insightful and minimal way and it will help to form the basis of the approach adopted here but there is a problem with Lombardi's approach which will mean that we cannot take it over without modification. I turn to this problem below, but first note some of the advantages of Lombardi's analysis.

While the approach is couched in a feature-geometric framework, the units assumed are entirely privative. The plain voiceless stop is characterised by the absence of any feature but

¹⁷ As we will see, Keating's (1984) conclusion is quite close to that which is drawn here, although there are crucial differences. Her distinction between phonetics and phonology is apt, but she seems to draw the line just in the wrong place.

this is not a case of underspecification; it is simply non-specification. The difference between non-specification and underspecification is that in underspecification, material which is vital to the phonetic interpretation of a segment is missing (and must often be added by arbitrary fill-in processes) whereas in non-specification this is not the case. Languages have the option to specify a segment with a laryngeal feature/element but this is not essential for the segment's interpretation. If no laryngeal activity is specified, none occurs but the segment can still be heard thanks to the other aspects of articulation.

One other notable aspect of Lombardi's model is the co-occurrence of more than one laryngeal feature in a segment. This allows for a highly economical system of privative features and is a principle adopted here in our reinterpretation of 'feature' as 'element'. We will see parallels for this behaviour in other types of element in section 3.1.5. The reinterpretation of such features as elements is not so straightforward as simply renaming them, however. Issues arise as to the precise properties that such elements might have and how we can expect them to behave phonologically. We can also consider to what degree they are interpretable on their own. I investigate these points later in this chapter.

There are certain problems with the way that we have been considering the question of laryngeal features up till now and there is a problem with simply taking over Lombardi's model entirely. One general point that I have been glossing over until now is that phonological segments do not necessarily remain the same throughout the phonology of a language. Phonological processes of the type that we have discussed in several places in this thesis (and will discuss again) can alter the underlying laryngeal specification, just as they can other aspects of the segment. For example, there is a well known process in English (and German and other languages) which accounts for the fact that stops which are typically described as underlying /p, t, k/ are sometimes [p^h, t^h, k^h] on the surface, and sometimes [p, t, k]. This is generally referred to as a process of aspiration (see, for example, Kahn 1976, Kiparsky 1979). This is naturally not something that we can simply ignore, and we need to consider carefully which of the two is the direct correspondent of the underlying segment and which is derived when such alternations are observed.

The key problem that we will encounter with the picture of laryngeal specifications in Lombardi (1991) is shared with many other approaches. It does not lie so much with the insight behind the set of features and their identification with abstract segment-types. The problem arises in the way that Lombardi applies some of her analyses in the languages which are most central to this thesis. It is this point that will drive much of the remaining discussion

in this section, although there are other points to be made. To the extent that her position is a reflection of a standard tradition, Lombardi can stand as an example of the problem.

We have identified the possible set of laryngeal contrasts in segments but we have not identified which of those segments occur in the languages that are the central source of data for this thesis. We noted previously that the 'starting point' for a serious investigation of the issues was that the two series of stops in German and English, and hence most likely their immediate ancestors, are distinguished by a binary opposition [\pm voice]. Now, Lombardi cannot use this kind of distinction because she sensibly uses privative features, but what is the distinction? While she does not focus at length on the phonology of the languages that concern us here, she does deal with German in some detail. One of her analyses, which accounts for Final Obstruent Devoicing¹⁸ reads thus: "German is a language that has the Voice Constraint: only onset consonants can bear [voice]" (Lombardi 1991, 49). While this has the effect of devoicing them in codas, this shows that segments such as /b, d, g/ are considered to be characterised by [voice] initially and presumably in underlying forms.

The problem, simply put, is that the stops of German are not voiced in initial position (nor are they elsewhere unless surrounded by voiced segments). Nor is there spontaneous voicing in the stops of English, nor those of most modern Germanic languages.¹⁹ One aspect of laryngeal phonology that does clearly characterise German, however, (and English and most other Germanic languages) is the presence of aspiration in the 'voiceless' stops. Both of these facts give these Germanic obstruents a very different kind of behaviour and nature to the obstruents of various other languages, such as typical examples of Romance and Slavic. To express this another way: in typical Romance and Slavic languages, such as Spanish and Russian, the stops in the series which are transcribed as /b, d, g/ surface as fully voiced (that is, with vocal fold vibration) in the majority of phonological environments (except where a clearly formulable process removes voicing phonologically); in these languages, stops in the series /p, t, k/ surface without vocal cord vibration of their own and without perceivable aspiration.²⁰ This is then different to the stops of English and German, where stops in the series which is typically transcribed /b, d, g/ do not surface as fully voiced, and /p, t, k/ are

¹⁸ See Brockhaus (1995a) for a detailed discussion of the issues surrounding Final Obstruent Devoicing in German.

¹⁹ The exceptions are normally claimed to be Dutch, Afrikaans and Yiddish, and we will discuss the issues that this raises below. It is also, in fact, too simplistic to see either German or English as a coherent unit in this regard. As we will also see later, different dialects, unsurprisingly, have different phonologies.

²⁰ I return to what is meant by the term 'aspiration' below.

clearly aspirated in most salient environment (unless a clearly formulable phonological situation prevents it).

This insight has been recognised repeatedly in Germanic phonology and those acquainted with it. There are classic statements of it in Sievers (1876, 1901) and Kuryłowicz (1948) and some recent statements of it are Kohler (eg, 1984), Goblirsch (eg, 1994). A form of it was also recognised by Jakobson (eg, Jakobson & Waugh 1979).²¹ Most approaches to the issues have not been tied to phonological theory, however, and so have not drawn all the possible conclusions. I argue below that they have also not yet been properly considered for the underlying level of phonology and have been overcomplicated by missing the effect of dialectal diversity. Importantly, recent theoretically-informed approaches have grappled with the issues and have come close to recognising the full implications of them. In the next subsection, I briefly review previous proposals and thinking on the issue and then go on to illustrate the implications.

3.1.3.1.4 Laryngeal Realism 1: *fortis/lenis*

The tradition in much writing on Germanic languages is to describe the distinction between stops as being a distinction between *fortis* and *lenis* series. This terminology developed in the philological tradition. There are various formulations of it and we will not investigate them all, but the crucial insight, expressed in terms of the Indo-European languages that European philologists have focused on, is that among those languages with two series of stops, different kinds of laryngeal distinction are made. This means in our terms that there are different kinds of segments in the languages. In what follows, I call this position ‘laryngeal realism’ because it seeks to represent the situation in all languages correctly and ties in with some of the notions explored for the historical dimension in section 2.3.

Much of the discussion of the *fortis/lenis* distinction has taken place in the ‘Germanic’ philological (and phonetic) literature. It has focused on the comparison of European languages from a philological perspective and has largely been satisfied to simply note the distinction.

²¹ Given Jakobson’s brilliant influence on generative theories of features (see, for example, Anderson 1985), it is interesting to see how this has almost fed into mainstream approaches, but not quite. As we shall see below, Jakobson suggests the use of a feature [\pm tense], rather than [\pm voice], in certain circumstances. The situation in Chomsky & Halle (1968) is not entirely clear in that they do use a feature [\pm tense] and mention a possible role in the make-up of consonants, but [\pm tense] is only used in rules which affect vowels and the obstruent series are distinguished by [\pm voice].

Given the universalist and uniformitarian position adopted here, however, we need to consider how the simple facts of German, English, Spanish, Russian and other well studied languages fit in with the set of possible laryngeal specifications that were outlined above. Two largely unconnected strands of research in phonology have recently converged to consider these issues, with remarkable agreement in the results. There has also been recent important work to support the phonetic foundation of the basic position, which also considers phonological issues and arrives at results that are compatible with the synthesis position that I will come to below. Firstly, I consider the insights and drawbacks of work in the fortis/lenis tradition.

We should return to some of the work cited in section 2.1.4, which describes the laryngeal states of segments in Present-Day German dialects. The presentation in that section follows their transcriptions precisely but, to give them their due, work such as Schübel (1955), Bock (1965) and Weldner (1991) recognise a fortis/lenis distinction, and describe the result of the IGCW as lenis. Works such as Kohler (1984) and Braun (1988) consider the fortis/lenis distinction in some detail, and Alexander (1983) considers certain diachronic implications and evidence from Germanic languages. Many of the results I derive from the final position that I adopt will be compatible with this work, but they are different in kind and in implication because they are clearly designed to fit with the work of more specialist laryngeal phonology, such as that in Lombardi (1991).

A possible synthesis of the fortis/lenis model of laryngeal specifications is this: the distinction between the two series of stops in languages like Spanish, French and Russian can be captured by the feature [\pm voice] (or maybe presence or absence of privative [voice]), whereas the distinction in languages like German, English, Danish and Icelandic can be captured by the use of [fortis/lenis]. Most of the key problems with the fortis/lenis position derive from the fact that, even though it seeks to be more specific than those who would have a simplistic voice/voiceless (ie, [\pm voice]) distinction in all languages with two series of stops, the feature distinction (or is it feature pair?) is often not clearly formulated. It is not clear that the fortis/lenis position is considered by its advocates to be privative (Kohler 1984 explicitly adopts [\pm fortis], for example) and it does not fit easily with any model of features. The distinction is sometimes seen as a 'special' mixture of voicing, aspiration and length, and this makes it incompatible with the minimalist assumption that features (or, rather, elements, once the translation is effected) are likely to have the same kind of status as each other.

Another key problem with the fortis/lenis model is that it does not go far enough in recognising the distinctions between languages. The fortis/lenis insight is based on the

recognition that there is a distinction between, say, German and Spanish in terms of laryngeal specifications, it often does not adequately take into account the fact that 'German' and 'Spanish' are problematic ideas ontologically. Part of the discussion in chapter 1 was devoted to the widely recognised idea that the modern socio-political entities such as 'German' are of only limited use in phonology and in historical phonology all the more so.

Phonetic and dialectological traditions have clearly shown that there is variety in terms of the laryngeal specifications of the linguistic systems that are often described to be 'dialects of German' and 'dialects of English' (see, for example, Goblirsch 1994 for a discussion of the situation in German and Harris 1994 for some of the details in English). As we saw in chapter 1, we do not need to abandon the socio-political notion 'non-standard variety' but we need to recognise the relation that it has to 'psychological real linguistic mental state'. In what follows, I argue that the fact that certain 'varieties of English' lack aspiration and that there are substantial differences in stops in 'varieties of German', for example, results from a difference in both underlying and surface laryngeal specifications. This simply becomes another of the phonological differences that exists between varieties and so, if we take 'English', say, to be a socio-political entity (an 'E-language' in the sense of Chomsky 1986), then we will not be able to speak of 'the laryngeal specifications of English', just as we cannot speak of 'the segmental inventory of English' in the singular because only some varieties have /ʌ/ or /θ/ or /y/, for example. We will, of course, be able to speak of the laryngeal system of any individual variety, including the 'standard' or 'reference' varieties such as RP and General American for English and *Hochlautung* (or *Bühnenaussprache*) for German, which naturally have an important role, but are too often the only varieties considered.

After a detailed survey of laryngeal effects in varieties of German, Goblirsch (1994) employs a *fortis/lenis* distinction to characterise all of the varieties, even though, as he explains, this groups together such different effects as aspiration, voice and segmental length (the first two involve very different action in the glottis, the last involves utterly different articulatory effects). This is not an approach that we can entertain here because it conflicts with the idea that subsegmental elements should be mappable onto unitary phonetic concepts and in any case, it seems unnecessary: from the perspective of historical phonology, there is no reason why all 'varieties of German', for example, should necessarily have innovated the same processes, and hence now have the same synchronic phonology. The very reason behind the existence of modern varieties of languages, indeed the very reason behind the fact that

there are different 'Indo-European languages', is that they do not. In fact, we shall see that the IGCW and the EIFV, as discussed in chapter 2, are examples of where the laryngeal specifications of some 'varieties of languages' have innovated a process, whereas others have not. It is processes like these that have produced the variation that now exists in E-languages.

Given this, it is impossible to speak literally of a 'Germanic system of laryngeal specifications' because there is variation not only between reference and non-reference varieties of individual E-languages, but also because there is variation among the reference varieties of those individual E-languages which have a Germanic heritage. As is frequently noted in the literature (eg, Cohen, Ebeling, Fokkema, van Holk 1972, Iverson & Salmons 1995, Lass 1997), the reference form of languages with origins in and connections to what is now the Netherlands (Frisian, Dutch and Afrikaans) do not feature aspiration in the stops /p, t, k/ and /b, d, g/ are typically fully voiced. This implies that they have a different laryngeal phonology; as has been remarked elsewhere, and as we will see below, this will turn out to be the same as that of Spanish, French and Russian.

Of course, it would be possible to speak of a 'Germanic tendency' in terms of laryngeal phonology, if the majority of Germanic linguistic systems turn out to function in the same way. It is also perfectly possible to speak of the 'laryngeal specifications of Germanic', when the term is taken to mean the psychologically real I-linguistic system that existed in the minds of speakers several millennia ago. In this usage, the word 'Germanic' is often prefaced with 'Proto-'.

The line of 'laryngeal realism' reasoning found in the fortis/lenis tradition, is that the obstruents of the reference varieties of German and English (and Danish, Icelandic, Norwegian and Swedish) are different from those of the reference forms of for example, Spanish, French, Russian and, it might be added, Japanese and many other languages. In the next section, following others, I build this into the laryngeal phonology of these languages. One potential barrier to the acceptance of this idea is the question of whether it truly fits the phonological facts. Keating (1984) argues, as quoted in section 3.1.3.1.3, that we should not take the phonetic evidence as sufficient for the postulation of such a distinction in phonology. This is quite right, but it is necessary to create an account which is compatible with the phonetic values of surface forms. In the next section, we will see that there is also phonological evidence for the distinction, but before we consider that, it will be well to be certain that the phonetic observations are on a strong footing. It would be otiose to examine every language mentioned above and below in this thesis, but we can quite easily take over the results of one detailed investigation of one of the languages which is most crucial here: German.

The most detailed and recent investigation into the phonetic properties of the obstruents of German is Jessen (1997, 1999). Jessen reports on the results of instrumental investigation into the acoustics and articulation of the obstruents found in a group of speakers who use what he calls ‘standard German’ pronunciation.²² We cannot discuss all of Jessen’s findings or methods, but his conclusion, after extensive demonstration, is unequivocal. In terms of the two series of stops in German, he shows that voicing is simply not the basis of the contrast. Aspiration, on the other hand, is very robust in key environments: “the mentioned groups of stops do not invariantly differ in voicing, since there are several contexts in which both groups are realized without voicing [...] The phonemes /p,t,k/ and /b,d,g/ ... were shown to differ in aspiration in all of the three representative contexts” (Jessen 1999, 299).

This agrees with reports for reference varieties of English and Danish, but contrasts with the reports for other languages and indicates that the phonetic basis for a phonological distinction is firm. This, and the other reasoning and evidence presented in this section from the fortis/lenis tradition seems to indicate quite securely that there is a ‘laryngeal realist’ distinction between the kinds of laryngeal specifications found in (many varieties of) German and English and those found in (many varieties of) Romance and Slavic languages. It seems that while Lombardi’s (1991) features may form a sound basis for an understanding of laryngeal phonology, they need to be combined with the fortis/lenis tradition’s recognition of differences among the laryngeal states that are found in individual languages.

The work reported in this section is important, but its impact is lessened by the fact that they do not situate the fortis/lenis distinction in a truly universalist phonology. From the perspective adopted here, the basic assumptions of many of the approaches are also problematic because they conflict with the minimalist nature that was predicted for phonology in section 3.1.1.

3.1.3.1.5 Laryngeal Realism 2: universalist

Alongside his investigation of the phonetic facts, Jessen (1997, 1999) considers phonological issues. The precise theoretical conclusions that he draws are not quite those that we will come to in this section, but they are compatible with the general thrust of what we will assume. Jessen revisits Jakobson’s (eg, Jakobson, Fant & Halle 1952 and Jakobson & Waugh 1979)

²² The definition of ‘standard’ here is problematic, as was discussed in chapter 1, but it is based on a reality of sorts which is spoken by a range of speakers; it is described as *Standardlautung* in the key authority, the *Duden Pronouncing Dictionary* (Mangold 1990).

system of features and characterises the distinction in German using [\pm tense], rather than [\pm voice]. While we would reject the use of binary features (along with Lombardi and many others for the laryngeal domain at least), the important point here is that standard German is given a completely different laryngeal specification to that which would be given for Spanish, for example. Two other strands of work in theoretical phonology have recently sought to do justice to the intricacies of laryngeal phonology. This work will allow us to capture the generalisations that have been discussed using elemental units.

The two strands occur in Harris (1994) and Iverson & Salmons (1995, 1999a). Both Harris and Iverson & Salmons work with privative units, so their proposals are eminently translatable. Harris (1994) is working in an explicitly elemental framework and he uses elements which he names [H] and [L]²³ which he glosses as ‘stiff vocal cords’ and ‘slack vocal cords’ respectively. Iverson & Salmons use [spread glottis] and [voice], which function in exactly the same way and, while they might predict slightly different actions in the glottis, we can take them as being equivalent.

Both Harris and Iverson & Salmons recognise that three types of underlying stop segment occur in the phonology of the two groups of languages identified in the last section. We can summarise the three types of segment are: (i) ‘fully voiced’ (ie, with vocal cord vibration throughout the segment), (ii) ‘voiceless unaspirated’ or ‘neutral’ (ie, with no vocal cord vibration and, if followed by a vowel, the voicing associated with the vowel commences immediately on release of the closure) and (iii) ‘voiceless aspirated’ (ie, with controlled vocal cord activity inbetween the fully open position for rest and the position for voice which causes a burst of ‘aspiration’ noise on release of the closure and leads to delay in the onset of voicing in a following segment which is specified for voice).

The key claim of Iverson & Salmons and Harris is that the contrast in the series of stops in phonological systems like that of standard Spanish, Russian and Dutch is (i) : (ii), whereas the contrast in linguistic systems such as the reference varieties of German, English and Danish is (ii) : (iii). As mentioned above, I call this insight ‘laryngeal realism’.²⁴

²³ In Harris’ exposition, elements are not encased in vertical slashes, but I impose on them my convention from the discussion of manner and place. These elements were first used in Kaye, Lowenstamm & Vergnaud (1989, 1990) but those authors do not discuss the facts of laryngeal realism.

²⁴ Ewen & van der Hulst (2001) show that this proposal can be incorporated into the system of laryngeal specification used in classical Dependency Phonology, which we have not had space to investigate here, but which has not been based on ‘laryngeal realist’ analyses. The system is intricate and would suffer from a short summary here, but one point to note is that aspiration is characterised by the use of [O], which can function as [H] or [spread glottis] do here.

One further important piece of evidence brought in to support this type of analysis comes from a recognition of which features can be seen to be active in the phonology of the two groups of languages. Iverson & Salmons (1995 and especially 1999) show that, while there is assimilation to [voice] in obstruents in languages like Dutch and Russian (in Dutch, for example, underlying /p, t, k/ surface as [b, d, g] in *klaphant* [klabbant] 'flat tyre', *potdicht* [paddixt] 'tight', *kookboek* [kogbuk] 'cookbook' - see Booij 1995),²⁵ no such process occurs in languages like English and German. On the contrary, both (standard forms of) these languages have processes which is often viewed as 'assimilation to voicelessness'. Such processes include sonorant devoicing (where a sonorant, which we would expect to be underlyingly voiced, assimilates to an adjacent obstruent, for example Northern English *plan* [pʌn], *treat* [tɹi:t], *crud* [kɹʊd], and Standardlautung German *Platz* [pʌts] 'place', *kriechen* [kɹi:çən] 'creep' - see Kohler 1985) and English plural formation (where an underlying /z/ - which surfaces unaltered after eg /b, d, g/ and vowels - surfaces as [s] in, for example Northern English *cats* [kʰats], *cups* [kʰʊps]).

Iverson & Salmons analyse this difference in the types of assimilation which can occur in languages as another facet of the laryngeal realist position. Those languages which feature [voice] in stops can have assimilation of the Dutch type, where the feature spreads to neighbouring segments, whereas those which feature [spread glottis] can have assimilation of the English type with [spread] spreading.²⁶

Iverson & Salmons (1995, 1999a) also show how a third feature, [constricted glottis], can function in stops in other linguistic systems, bringing the number of features which they assume to three, which is analogous to the system assumed by Lombardi (1991). In fact, Lombardi's system and that used by Iverson & Salmons are easily translatable. The equivalences are shown in (3.7); Harris' elements are also included here for comparison, as is one interpretation of Jessen's (1997, 1999) Jakobsonian features:

²⁵ Some further aspects of the laryngeal phonology of Dutch are dealt with in section 4.2.2.

²⁶ This follows simply for 'sonorant devoicing' if [spread] is assumed to equate to voicelessness, which seems reasonable given the effect predicted for the vocal cords. This basic analysis is also assumed in Anderson & Ewen (1987), where [spread glottis] is equatable with their [O].

(3.7)

Lombardi	Harris	Iverson & Salmons	Jessen
[aspiration]	H	[spread glottis]	[+tense]
[voice]	L	[voice]	[+voice]
[glottalization]		[constricted glottis]	

While there are differences of details in the implementation of the proposals, and Lombardi does not recognise laryngeal realism, the three proposals can be interpreted along the same lines. Iverson & Salmons (1995) show how their features can characterise all the segment types given in (3.4) and this allows them to account for all possible laryngeal systems. They use a similar idea to Lombardi and allow the features to co-occur to produce laryngeal specifications for segments such as /b^h/ (familiar to us from in the traditional reconstruction of Indo-European) using a combination of [spread glottis] and [voice]. Ejectives are characterised as featuring [constricted glottis], and voiceless aspirated segments such as /p^h, t^h, k^h/, in languages like Hindi, which has four contrastive series of stops (eg, at the labial place of articulation: /p, b, p^h, b^h/), are characterised as underlying [spread glottis]. All of these features are used to the full in phonology and all possible combinations are attested. The proposals considered here link the universalist concerns of theoretical phonology with the philologically recognised need for laryngeal realism.

3.1.3.1.6 What elements are needed for laryngeal specifications? Take 2

The task of proposing a set of elements that will be used in further analyses in this thesis is now largely a terminological matter. As we have seen in all of this subsection (3.1.3.1) up till now, proposals exist which capture the facts of laryngeal phonology excellently and the units that have been suggested in these proposals typically have many of the properties ascribed to elements at the start of this wider section (3.1.3); this has been overtly recognised by some theorists, and it is a simple task to set out an element set in (3.8). Following the practice developed for elements of place and manner, I name the elements according to the main property that they relate to in articulation, so that they are broadly interpretable. (3.8) also includes a brief indication of the types of segment which include the element in their make-up:

(3.8)

|spread|

- in aspirated segments

|voice|

- in fully voiced segments

|constricted|

- in 'glottalised' segments, such as ejectives and implosives

The names given to the elements are basically those of Iverson & Salmons and I take over their characterisation of stop laryngeal specifications. If we assume the position of laryngeal realism, then, for a language with two series of stops, we can speak of (i) '|spread| languages' where |spread| characterises one series and the other is the non-specified 'neutral' segment and where |spread| can be expected to be active in the phonology in spreading; (ii) '|voice| languages' where one stop series will feature |voice| and the other will be 'neutral' and where |voice| is expected to be active; (iii) '|constricted| languages' also occur, although we will not discuss any here (Maddieson 1984 lists six - ejectives are more common as a third series).²⁷ Note that these designations refer to the series of stops; we turn to fricatives later (in section 3.1.3.1.11). On a phonological level, we might expect the neutral series in both |spread| and |voice| languages to be the unmarked series because it is a phonologically simple entity. The names given to the elements do not necessarily correlate with the way that their presence is observed; it is clear from phonetic investigations that they can affect the length of adjacent vowels, for example.

Given their status as elements, there are some further questions to be considered. One has to do with the question of 'stand-alone interpretability', ie, whether the elements can be interpreted as segments in their own right. I consider this in section 3.1.3.1.8. A further point has to do with the precise status of aspiration in Germanic and other languages (investigated in section 3.1.3.1.9). A final question, which will be important given the data in this thesis, is whether these elements can occur in fricatives just as in stops; I turn to this in section 3.1.3.1.10.

²⁷ We might note here that Keating (1984) argues against the position adopted here (and see Wagner 2000 for further discussion) but it is remarkable how similar some of her conclusions are. She recognises that there are three fundamental phonetic types of segments: {voiced} {voiceless unaspirated} and {voiceless aspirated}; however, she goes on to argue that all these three are derived from an underlying abstract distinction between [+voice] and [-voice] which does not relate simply to these observable entities. It seems a more minimalist position to assume that the three categories, which are necessary for description at one level, are all that occurs throughout phonology.

Before we examine these points, however, it will be worth exploring the effects of laryngeal realism on the type of data which is central to this thesis. The processes presented in chapter 2 were described in terms of segmental units, as is normal in the philological literature. The key reasoning behind the introduction of phonological theory into this thesis is so that the processes involved may become more comprehensible and chapter 4 engages in the process of phonological analyses of the GCS, HGCS, EIFV and IGCW. For this reason we will not consider the elemental make up of the segments involved there in this chapter but it will be clear that the position of laryngeal realism adopted here will have quite an effect. The main problem is that the symbols used in such descriptions are inadequate, a point made by Harris (1994).

We saw in section 3.1.3.1.1 that the basic symbols which are at the disposal of the phonologist derive originally from a language with only two series of stops. It seems likely on comparative grounds that Latin was a [voice] language, given that Present-Day Romance languages are overwhelmingly [voice] languages. It is straightforward to apply the two series of letters developed to write a [voice] language to write the segments in a [spread] language, and because writing does not require phonological or phonetic preciseness (merely an indication of contrastiveness is helpful), it is entirely reasonable to do so. However, it is not so reasonable to have ambiguous phonetic or phonological symbols. If laryngeal realism is correct, and as we have seen, the evidence is compelling that it is, then we must conclude that the symbols /p, t, k/, for example, have been used to transcribe two different kinds of phonological object (both underlyingly 'neutral' and underlyingly 'aspirated' segments). This need not matter if they are simply used in the transcription of one language at a time and the value that they are given is made explicit, but when comparisons between systems are to be made and when universal claims as to inventories or the nature of possible phonological processes are to be made, the ambiguity is extremely unhelpful. It may even be that it could mislead the analyst into grouping together segments, inventories and processes which do not belong together. It is becoming apparent that new symbols are needed for segments, and we turn to this in the next section.

3.1.3.1.7 Symbols for segments

In section 3.1.2, quite some space was devoted to the justification of the use of the segment in phonology. The foundation of the current section (3.1.3) is that segments are composed of

elements, but just as the existence of sub-atomic particles does not invalidate the concept of the atom, so it is with phonological elements and segments.

It will be helpful, in the pursuit of clarity in exposition, to give segments symbols that are unambiguous. This means that we will need to make a distinction between the symbols used to transcribe the obstruents of |spread| languages and those of |voice| languages, just as we would for those of |constricted| languages. Naturally the IPA gives us the resources to transcribe all possible segments, but there are not enough base symbols for all the types of obstruents, so it will be necessary to use diacritics. This is already the case for |constricted| languages, where it is conventional to use /pʰ, tʰ, kʰ/ for ejectives, for example.

A symbol is proposed in (3.9) for each of the three types of segment which occur underlyingly in languages like English, German, Spanish and Dutch. Laryngeal elements are indicated in (3.9) for those segments which feature such an element. The exemplification features stops and fricatives, at the classic three places of articulation, one for each of the elements |labiality|, |coronality| and |dorsality|; I discuss the issues raised by the extension of the reasoning for laryngeal realism from stops to fricatives in section 3.1.3.1.10, but the logical possibility of the segments is clear:

(3.9)

‘neutral’	‘voiceless aspirated’ spread	‘voiced’ voice
p ^o	p ^h	b
t ^o	t ^h	d
k ^o	k ^h	g
f ^o	f ^h	v
s ^o	s ^h	z
x ^o	x ^h	ɣ

The symbols used here are not exceptional. No heuristic diacritic is used for /b, d, g/, which are given their normal IPA value of ‘voiced’, but from now on, these symbols are used unambiguously for segments which feature the element |voice| in their make up. The symbols /p^h, t^h, k^h/ hardly need any explanation. They are conventional in square brackets for most Germanic languages and are in use for underlying segments in languages like Hindi, Thai and Korean where there are uncontroversially underlying voiceless aspirates as one of three of

four series of stops and also in languages like Azerbaijani, which has only two series of stops which are typically transcribed (eg, in Maddieson 1984) as a contrast between /p, t, k/ and /p^h, t^h, k^h/.

The use of the diacritic ‘^o’ diverges from recommended IPA (see, IPA 1999) and it is chiefly employed here for clarity. It is reminiscent of the IPA ‘devoiced’ diacritic [̥], but is simpler typographically and is used to remind the reader that the segment is not underlyingly specified for laryngeal activity (so it does not have the dynamic implication that might be read into the ‘devoiced’ diacritic). Standard /p^o, t^o, k^o, f^o, s^o, x^o/ are used for ejectives.

One obvious implication of this is that certain types of phonological segment have been represented in spelling in more than one way, and conversely one set of letters (eg <b, d, g>) has been used to represent more than one kind of phonological segment. To provide for absolute clarity, the correspondences for (|spread| varieties of) English and (the |voice|-language) Spanish are given in (3.10):

(3.10)	Letters	English	Spanish
	<p>	/p ^h /	/p ^o /
	<t>	/t ^h /	/t ^o /
	<k>	/k ^h /	/k ^o /
		/p ^o /	/b/
	<d>	/t ^o /	/d/
	<g>	/k ^o /	/g/

This results from the situation that was described in section 3.1.3.1.1. The alphabet that was available and was the obvious choice for English and Spanish when they first were written in the scribal tradition, which we basically continue today, was the Roman alphabet, and this had two ‘series’ of letters for the representation of stops; the choice for both languages was obvious and we have not changed these traditions as they serve us perfectly well. It will be helpful at times to have a way of describing this situation; to do this I readapt the terminology used for classical descriptions of IE consonants. I adopt the following terminological conventions from now on: (i) *Ts* (from ‘Tenues’) = phonological segments which are represented orthographically with the letters <p, t, k>, (ii) *Ms* (from ‘Mediae’) = phonological segments which are represented orthographically with the letters <b, d, g>.

Section 2.3 was devoted to the reinterpretation of the phonological processes of the GCS, the HGCS, the EIFV and the IGCW in accordance with standard IPA conventions and

phonological concerns. The import of laryngeal realism for these processes will be examined in chapter 4, but it will be clear from this subsection that part of that process will involve a symbolic reinterpretation. The remaining subsections in this section address further points which will support the analyses adopted here.

3.1.3.1.8 Laryngeal elements alone: debuccalisation

We saw at the start of this section (3.1.3) that the basic vocalic elements |I|, |A| and |U| are interpretable by themselves (as /i/, /a/ and /u/). This shows that segments can consist of just one element and, conversely, elements can in principle form segments by themselves. The possibility is thus open that laryngeal elements can be interpreted alone as segments.²⁸ In this section I investigate certain aspects of this point which are in their own right important aspects of the interpretation of the elements, but which will also be important evidence for the set of elements assumed here.

It was further claimed at the start of this section that phonological processes can cause elements to be lost from segments and this opens up the possibility of investigating whether, and if so, precisely how, the laryngeal elements can be interpreted as segments by themselves, because, if an obstruent consists of place, manner and laryngeal elements and the place and manner elements are lost, this will leave only the laryngeal elements, and if this results in a segment, that will be the stand alone interpretation of the element concerned. The kind of argumentation used here comes from a tradition which recognises debuccalisation (or 'deoralisation', ie, the loss of supralaryngeal articulation while retaining laryngeal articulation) as a frequent phonological event; it goes back to Lass (1976) and has been developed in the Dependency Phonology and Government Phonology literature, but also features in feature geometric work, such as McCarthy (1988), Keyser & Stevens (1994) and Fallon (1998).

In section 2.1.1.5, we saw that the fricative output of GCS₁ soon debuccalised in Germanic. The two stages of the processes involved were given in (2.7) and are repeated here as (3.11); this is altered to be in line with the comments in section 2.3.1 but not yet with laryngeal realism:

²⁸ It may be that not all elements have this property. This is as yet uncertain and requires future research. It may be that 'manner' elements require a 'place' element in order to be interpretable.

(3.11)

$k > x > h$

As was pointed out previously, this is a classic type of lenition trajectory and such processes are common (see Dosuna 1996) where fricatives are the input and the glottal fricative is the output. Now, the segments in (3.11) have not been analysed in terms of laryngeal realism, but it is at least possible that the make up of the fricative contains |spread| (it must either contain this or have no laryngeal specification)²⁹ and so this is a candidate process which could reveal the stand alone interpretation of that element. If we consider the fricative segment in (3.11), the only other elements in its make up are |dorsality| and |frication|. ³⁰ It is clear that the element left after debuccalisation is not |dorsality|, because it is precisely oral information which is lost. The other candidate is |frication|. It has been proposed in Harris (1990) that the segment /h/ does indeed consist of only the element |frication| (or at least of his version of the element, which he labels |h|) and the evidence of the debuccalisation process in (3.11) is not sufficient to decide the issue. However, there is other evidence which I believe shows that the debuccalisation product of segments, such as the dorsal fricative in Germanic, is |spread| which is then interpretable as the only element in the segment /h/.

This evidence comes from a process which can be seen almost as the ‘opposite’ of this kind of debuccalisation. Fallon (1998) calls this process ‘fusion’ and traces it back to such work as Stahlke (1976), Schane (1984) and McCarthy (1989). In fusion, two segments coalesce (hence it is also referred to as ‘coalescence’, for example in Kiparsky 1988) and the product is one segment which features the elements of both. We have already seen a case of fusion, although it was not described in that way when we first encountered it (in section 2.3.2). The process was what lay behind the formation of the ‘TA’ segments in Indo-Iranian and was given in (2.32). I repeat it here as (3.12):

(3.12) Indo-Iranian TA formation

$ph > p^h$

$th > t^h$

$kh > k^h$

²⁹ Evidence will be presented below that the Germanic fricative was indeed /x^h/.

³⁰ This anticipates the summary on which elements occur in which segments in section 3.1.5 somewhat, but only a little and to the extent which is necessary. Further details are given there. This section also anticipates the discussion of lenition processes in section 3.2.2, as will become apparent there.

In this process the IE laryngeal segments fused with the IE *T*s, to give segments such as /p^h, t^h, k^h/. While we will not consider this in detail, it makes perfect sense to assume that the segments fit with laryngeal realism if the input consists of segments such as /t^o/ and the output of segments such as /t^h/ (as already transcribed). This will be further justified in later sections. These segments have come down unaltered into languages like Hindi and Iverson & Salmons (1995) show that they can be characterised in such Present-Day languages with the element [spread]. If we assume, with Kuryłowicz (1977) and Bammesberger (1989), that there was only one laryngeal in IE, then it would make a highly natural inventory if we assume that this laryngeal was /h/. In that case, Indo-Iranian *TA* formation is a fusion process which shows that a laryngeally neutral segment such as /t^o/ fuses with [spread] (that is, the segment /h/) to become /t^h/. If this is true then we must interpret debuccalisation to /h/ as debuccalisation to [spread] and this is evidence for the stand-alone interpretability of the laryngeal element.³¹

The evidence presented here is perhaps a little too infirm a foundation to base such solid conclusions on. But there is further evidence which is much more secure: Hock (1986) reports a diachronic case from Korean and Maza (2000a,b) reports a synchronic case in Granada Spanish. The basics of Hock's data is presented in (3.13):

(3.13)

manh-ta > mant^ha 'be much'

noh-ta > not^ha 'set free'

This shows that whatever it was that occurred originally in final position in the preceding morpheme has fused with a following *T* to produce /t^h/. Maza's (2000a,b) data shows a remarkably similar situation. Maza deals with aspects of the quite well known (see, for example, Dosuna 1996) process of the debuccalisation of /s/ in Andalusian Spanish. In this process underlying /s/ is realised as [h], producing such forms as *pozos* [poθoh] 'wells' and *libros* [liβroh] 'books'. Vaux (1998) describes a similar situation and shows that the original /s/ fricative is specified for [spread], so the debuccalisation leaves just the laryngeal element. However, Maza (2000a,b) shows that, where [h] occurs preceding a *T*, the two can fuse and the result is an aspirated voiceless stop. Some of her data is given in (3.14):

³¹ Note that if this analysis is accepted, it is also evidence for the interpretation of the laryngeal segment (or maybe of only one of the laryngeal segments) in Indo-European.

(3.14)

[ɛt ^h e]	<i>eSte</i>	‘this (masculine)’
[æ:k ^h o]	<i>aSco</i>	‘disgust’

It thus seems that there is good evidence that |spread|, which we have previously seen to be realised as aspiration in stops, is realised as the segment [h] by itself. We will not devote so much space to the interpretation of the other laryngeal elements because they are not as crucial for this thesis, but a short treatment is in order. Fallon (1998), in a monumental, highly insightful and almost unique study of the synchronic and diachronic phonology of ejective stops, discusses several cases of debuccalisation in ejectives which seem analogous to the cases of Germanic dorsal fricative debuccalisation (and Spanish sibilant debuccalisation) just discussed. Fallon adopts the same type of laryngeal specifications used here (using a privative feature [constricted glottis] for ejectives and [spread glottis] for aspirated stops). One example is taken from Menz Amharic, where the ejective /k'/ has debuccalised to produce /ʔ/. Some of Fallon's data (from Cowley et al (1976)) is shown in (3.15). The evidence compares Menz Amharic with other varieties (which still feature the original segment):

(3.15)

Other varieties	Menz Amharic	
[lək'en]	[ləʔen]	‘for a day’
[jik'rebu]	[jiʔrebu]	‘let them come closer’
[awwek'e]	[awweʔe]	‘he knew’

It seems from this and the other data that Fallon (1998) discusses that ejectives debuccalise to /ʔ/, so we may assume that the stand-alone interpretation of |constricted| is the glottal stop. Fallon also presents evidence from fusion which corroborates this.

It is not clear that |voice| can be interpreted as a segment by itself. While this seems to destroy the unity of the set, which might be problematic of minimalist grounds, it is likely that there are good phonetic reasons for this. A segment like /ʔ/ (which only consists of |constricted|) is clearly perceivable as an interruption in the phonetic signal. The segment /h/ (which only consists of |spread|) is clearly perceivable when adjacent to any voiced segment, including vowels, because it has the effect of ‘devoicing’ these segments (in a similar kind of process to what was described as sonorant devoicing in section 3.1.3.1.5); there are even

phonetic arguments for transcribing a sequence such as /hat/ as [ʔat]. An element like |voice| would only be perceived as voicing, however, and as a segment in the vicinity of any voiced segment, that simply could not be perceived. So it seems likely that |voice| has no independent pronunciation for functional reasons.

In investigating the evidence provided by debuccalisation and fusion, this section has shown that the interpretation of laryngeal elements that was proposed in section 3.1.3.1.6 is on the right lines. It seems that |spread| inheres in ‘voiceless aspirated stops’ (and can also occur in fricatives) and that its independent realisation is as /h/. The next section briefly revisits one of these points (aspiration in stops) to reinforce the analysis of laryngeal realism.

3.1.3.1.9 (De)aspiration and voicing in Germanic languages

The approach that we have assumed here claims that the *T*s of reference varieties of Germanic |spread| languages like standard German and English are, in effect, underlyingly aspirated. It also claims that the *M*s are underlyingly unvoiced in that they are neutral stops, with no laryngeal element. It is well known, however, that aspiration is only clearly audible in *T*s in certain environments, for example, word-initially and foot-initially and that there can be vocal cord vibration during the production of some *M*s. If the *T*s are underlyingly aspirated, then there must be a process of deaspiration in certain environments, for example, intervocalically. This does not seem particularly problematic because the alternative, standard analysis is faced with exactly the same (or rather, the opposite) situation: if the *T*s are underlyingly unaspirated, then there must be a process of aspiration; if the *M*s are underlyingly voiced, then there must be a process which removes the voicing in most environments (apart from in environments where the segment is surrounded by clearly voiced segments, such as vowels). I tackle these two points separately in this section, but the end result will be the same.

One argument in favour of the laryngeal realist analysis proposed here for the *T*s of |spread| languages is that it can be seen as ‘simpler’ or, at least, less arbitrary phonologically to lose something which is underlyingly specified than it is to have something added for which has absolutely no source in the surrounding phonological environment. It is also worth noting that aspiration in stops is not simply a matter of Voice Onset Delay (see, for example Lisker & Abranson 1964, who propose measurements of Voice Onset Time as an account for laryngeal specifications). While this is an acoustically measurable correlate of aspiration, it is not the

cause of aspiration. As, among others, Catford (1988) points out, aspiration is caused by a positive glottal articulatory gesture, not by the absence of any articulation.

Another argument, as Spencer (1996) points out, is that even in those environments which are typically described as featuring unaspirated *T*s, there often is at least weak aspiration, for example word-finally and foot-medially. In fact, Spencer (1996) concludes that aspiration is possible in all environments except following an /s/, so it seems that the overwhelmingly more common realisation is the aspirated allophone; this is compatible with Jessen's (1999) observations for German. These points speak in favour of an analysis of the languages in question as [spread] languages which feature a process which deaspirates the *T*s, at least to an extent, in certain phonological environments.

In this connection, a possible argument against the laryngeal realist position would arise if such deaspiration processes were never observed in languages which uncontroversially have a series of underlyingly aspirated (ie, [spread]) voiceless stops, such as Hindi and Thai. Happily, however, processes of this type do exist in such languages, as Houlihan (1977) and Vijayakrishnan (1999) report: for example, in the Hooghli dialect of Bangla (Vijayakrishnan cites Ghosh 1995 for this) where stops are deaspirated in environments which Vijayakrishnan equates entirely with those where there is a lack of aspiration in English *T*s. There seems to be no compelling reasons to reject laryngeal realism on these grounds.

If we turn to the 'voicing' of the *M*s in [spread] languages, it is clear from phonetic analyses, such as Jessen (1997, 1999), that these segments, which I transcribe as /p^o, t^o, k^o/, typically only show any evidence of vocal cord vibration when they are in an environment which sees them surrounded by segments which clearly have spontaneous voicing of their own, such as vowel and sonorants. It seems entirely reasonable to propose that the voicing shown on the *M*s of [spread] languages is purely a coarticulatory carry-over from the voiced segments onto the stops. It is no surprise at all that neutral segments, which have no laryngeal element should be susceptible to such coarticulation, and there is no need to view this as a phonological process of element-spreading. Neutral stops contain no instructions as to vocal cord activity at all, so we expect the kind of activity that has already commenced in the production of an utterance should not stop immediately where there is no need to; it is also perfectly possible that voicing may commence anticipatorily where a neutral stop is followed by a voiced segment. There is equally little reason to reject laryngeal realism on these grounds.

3.1.3.1.10 Laryngeal specifications in the segments of Proto-Germanic

Given the situation in contemporary Germanic languages, such as the reference varieties of German and English just discussed, we might wonder what the laryngeal specifications of Proto-Germanic were and whether these Present-Day languages have simply inherited Germanic specifications or whether they have innovated new ones. As may be obvious, this will be rather important in chapter 4, when we return in detail to the Germanic lenitions from chapter 2. Proto-Germanic is the mid-point between the GCS and the HGCS, EIFV and IGCW and it will be crucial to establish the laryngeal specifications of the language. As can be seen from the result of the GCS in section 2.1.1 and from a comparison of all Present-Day Germanic languages, Proto-Germanic had two series of stops. It is thus a candidate for the distinction between *|spread|* languages and *|voice|* languages.

We will return to general and specific questions concerning the laryngeal specifications of early Germanic languages in chapter 4 and we will need to consider there the situation in Indo-European, as well. In this section I focus purely on Proto-Germanic, partly to establish the degree to which it is reconstructible and partly because the treatment of this issues here will facilitate coming discussion.

Reconstruction of such aspects of a language's phonology is not entirely straightforward. This is at least in part because, as we have seen, it is not unambiguously recorded in early spelling. Indeed, this very fact, it was argued above, is one of the key reasons why laryngeal realism has often not been properly recognised. This point is more obviously relevant to the historic Germanic languages and not directly to Proto-Germanic because we have no written records for the Proto-Germanic language itself but it is relevant here, too, because the early written records of Germanic languages are the best evidence that we have for Proto-Germanic, as Proto-Germanic is largely reconstructed from them using the comparative method.

There are, though, two kinds of evidence that can be used to help solve the question. These are: (i) a comparison of Present-Day Germanic languages, which we can easily analyse on this issue and where we are not reliant on spelling and (ii) a consideration of the patterns of laryngeal assimilation found in the earliest texts; these may be recorded in spelling and could be seen in alternations in a morphological paradigm between symbols used for *T*s and those used for *M*s.

It has to be said that there is a general unanimity among those who have actually considered the problem (rather than those who have not recognised laryngeal realism and

simply assume that Germanic had a [±voice] distinction). When we compare existing West and North Germanic languages, we find that there is remarkable consistency among them in this regard. Apart from the already noted ‘Netherlands group’ (which is mentioned in section 3.1.3.1.4 and elsewhere here and which we will return to in section 4.2), Germanic languages (bearing in mind the usual caveats regarding the notion ‘language’) all have a distinction of aspiration, not voice, that is, they are all |spread| languages. This is particularly well recognised for Icelandic and Danish as the issues of aspiration in *T*s and voicelessness in *M*s has been investigated in some detail for these languages (see Jessen 1999 and the references given there). Alexander (1983) focuses on Proto-Germanic and the laryngeal question, bringing in a range of evidence from several languages, and concludes that the contrast in obstruents was based on fortis/lenis, not voicing. While we have rejected the notion of fortis/lenis as a phonological feature, the laryngeal realist perspective takes over much from the fortis/lenis idea, and we can recognise the basic result of Alexander’s consideration of the question. Finally in this regard, Lass (1984b) sums up the comparative evidence for Proto-Germanic as a |spread| language well (although he is working from a perspective that sees aspiration as a phonological rule): “I think a good case can be made for aspiration as a common Germanic rule, differentially lost in the dialects; certainly it is now widely distributed in the family (many forms of English, Icelandic, Danish, Norwegian, Swedish, many German dialects, some Yiddish)” (1984b, 287). The obvious comparative conclusion is that Proto-Germanic was a |spread| language. Iverson & Salmons (1995) assume this as well.

There is even more evidence for this conclusion. As we saw in section 3.1.3.1.5, the patterns that can be seen in laryngeal assimilation in |spread| languages are different to those in |voice| languages. This is predicted by the model, in fact: if a language only makes use of |spread| in stops, then we expect that only |spread| will be the active in the phonology of that language in this regard and the same naturally applies to |voice| languages *mutatis mutandis*. Once this is recognised as a differentiating characteristic of the two laryngeal types of language, it can be used as a diagnostic. If we apply this line of reasoning to Germanic, we might wonder if there is any evidence for such assimilations.

There is naturally not a great deal of such evidence because such allophonic processes are difficult to reconstruct as they are only very rarely noted in writing. In any case, we have no direct written records from Germanic. If we consider Gothic, however, which is the earliest recorded Germanic language and is thus of particular importance for the reconstruction of Germanic, we find that the only recorded laryngeal assimilation process (as reported by

Braune & Ebbinghaus 1981) is of the |spread| type. This is a processes which is part of a wider phenomenon, but can be shown to involve neutral fricatives undergoing laryngeal assimilation to the extent that the author (Wulfila) wrote them with the Gothic letters for <f, s, h>, not <b, d, g> (the latter have convincingly been shown to represent fricatives in medial position, see Marchand 1973, Braune & Ebbinghaus 1981). Evidence for the process is found in the morphophonological alternations shown in (3.16). The first column of (3.16) shows an infinitive with the unassimilated fricative, and the second column features a derived form which shows the assimilated segment, which has taken its laryngeal state from a following *T*. The third column gives an informal representation of the assimilation in a rule-like format; the key point here is that the following segment in all cases is a *T* which I have indicated as /t^h/ on the assumption that its |spread| element has spread to the preceding segment:

(3.16)

<i>giban</i>	‘give’	<i>fragifts</i>	‘conferment’	$v^o > f^h / __ t^h$
<i>anabiudan</i>	‘order’	<i>anabaust</i>	‘ordered’	$\delta^o > s^h / __ t^h$
<i>magan</i>	‘can’	<i>mahta</i>	‘could’	$\gamma^o > x^h / __ t^h$

The assimilation here does indeed seem to be is assimilation to |spread|, and there is no evidence at all for assimilation to |voice| in Gothic, which, as the oldest attest Germanic language, gives the best window on Proto-Germanic. In fact, assimilation to |voice| seems to be unknown in the Germanic languages (apart from in the special case of Dutch and allied languages). This is further evidence for the assumption that Proto-Germanic was a |spread| language

Taken together, the two types of evidence discussed in this section provide a firm foundation for the assumption that Proto-Germanic was a |spread| language, with the series /p^h, t^h, k^h/ : /p^o, t^o, k^o/. This is an interesting result in its own right, but it also begins to illustrate the applicability and implications of laryngeal realism, to which we return in detail below.

3.1.3.1.11 Laryngeal specifications in fricatives

In section 3.1.3.1.7, I simply assumed that the laryngeal specifications that are justified for stops are simply transferable to fricatives. This is a reasonable null hypothesis, in that we would expect some explanation if the elements that can occur in certain obstruents cannot

occur in others, but there may be such reasons and some consideration must be devoted to the topic.

As was noted earlier, the majority of work that has investigated laryngeal specifications has focused on their role in stops. A certain amount of work has been carried out on the laryngeal states of fricatives, however, and some of the most important work is considered here. One simple point to note is that ejective fricatives are clearly attested in the languages of the world (for example, Tlingit and Hausa, see Ladefoged & Maddieson 1996), so |constricted| can indeed occur in fricatives. The main question in this regard is: does ‘laryngeal realism’ apply to fricatives? Do all the types of fricatives given in (3.9), and adapted and extended in (3.17) occur?

(3.17)

‘neutral’	spread	voice
f ^o	f ^h	v
θ ^o	θ ^h	ð
s ^o	s ^h	z
ʃ ^o	ʃ ^h	ʒ
x ^o	x ^h	χ

One point to note here is that there are languages that contrast these three types of fricative on the surface, just as there are languages which contrast all three series of stops. Ladefoged & Maddieson (1996, 179) show that Burmese contrasts all three, as shown in their data (with my interpretation in the top line), given here in (3.18):

(3.18) Burmese laryngeal contrasts in fricatives

‘neutral’	spread	voice
sà ‘to be hungry’	s ^h à ‘letter’	zà ‘lace’

From this, we can take the fact that all three kinds of fricative can occur in languages: these segments would map straightforwardly onto our /s^o, s^h, z/. This means that one aspect of the understanding of laryngeal specifications developed here is confirmed for fricatives as well as

stops. The other aspect is the idea that there are |spread| languages and |voice| languages. While we have seen that there is good reason to accept this for stops, is there for fricatives?

If any idea of 'economy' of element use holds (along the lines of Martinet 1955), then we might expect a language which has a spread distinction for stops to have a |spread| distinction for fricatives, and the same would apply for |voice|. But this is not clearly the case and it is not necessarily precluded that a language could have different types of contrasts in stops and fricatives; this is potentially a fruitful source for future investigation.

It is true that there is no perceivable Voice Onset Delay in the fricatives of English or German, but as we saw in the last section, this is not thought of as the prime phonetic method of exponence of the element. The action of controlled glottal spreading, which largely is expected from the element, on the contrary, can occur in the production of a fricative just as it can for a stop. One key source of evidence for the possible nature of the underlying distinction in fricatives in a language (for both the analyst and the child acquiring the language) is the phonological activity that the segments engage in, as Iverson & Salmons have shown. On this reasoning, we can note that sonorant devoicing is triggered by the 'voiceless' series of fricatives in those varieties of English and German where it is also triggered by stops, and this could be precisely the evidence that a child needs to recognise a |spread| specification and to set up a series of fricatives such as /f^h, s^h, x^h/ against /f^o, s^o, x^o/.

Vaux (1998) has recently shown that languages where fricatives have previously been described as showing no evidence of |spread| (such as varieties of Armenian) do in fact have an active |spread| specification in one of the series. One implication of this is that there have been cases of mistranscription (or, perhaps, ambiguous transcription) in fricatives, just as there have been in stops. The main aim of Vaux (1998) is to show that 'voiceless' fricatives can be positively specified for |spread|. ³² This fits well with the picture in (3.16) in that it shows that fricatives do not have to show aspiration to be specified with |spread|. We can fairly assume that those fricatives that show no phonologically active feature can be seen as being laryngeally non-specified, and there seems to be no reason not to extend laryngeal realism so as to allow for languages which contrast two series of fricatives by using only |voice|. Furthermore, the evidence seems to show that Present-Day Germanic languages, such

³² As Vaux is working in a model with binary features, he actually expresses this as [+spread], but the point is the same as the privative formulation.

as reference varieties of English and German show the same distinction in fricatives as they do in stops.³³

3.1.3.1.12 Summary and prospect

This section has argued in some detail for both a set of laryngeal elements and for a way of applying them in languages. The detail was required because the conclusions are quite far-reaching. The fact that I had to propose a new set of unambiguous symbols in section 3.1.3.1.7 to enable clear discussion of the true situation in languages is a fair indication of the potential impact of the proposals. If the position of laryngeal realism which was recognised in this section is correct, then it clearly has substantial implications for our understanding of historical phonology, just as it does for contemporary synchronic phonology, and we began to explore some of the historical implications in section 3.1.3.1.10, where it was shown that Proto-Germanic was a *|spread|* language. In order to properly understand the four sets of changes described in chapter 2, we will need to reanalyse those, too, in terms of laryngeal realism. Doing this will not be entirely a simple matter, as the phonological system of the languages involved and of their daughter languages will need to be considered.

After a short next section on prosody, section 3.1.5 puts the elements proposed here together with those proposed earlier for place and manner to illustrate the representations that are assumed for whole segments. It will be clear there, just as it is here, that the representations that we adopt will have quite some implications for the simple types of lenition trajectories discussed in section 1.4. For example, an equation such as ‘ $t \Rightarrow d$ ’ needs now to be reinterpreted. Does it mean $t^{\circ} \Rightarrow d$, $t^h \Rightarrow t^{\circ}$, or $t^h \Rightarrow d$? Are all these changes lenitions? Are they all, in fact, possible changes? Equally, the change ‘ $p \Rightarrow f$ ’ becomes problematised: can both $/p^h/$ and $/p^{\circ}/$ undergo this process? And is the output $/f^h/$ or $/f^{\circ}/$?

These questions are investigated in section 3.2 when we turn to a general consideration of lenition. Any change in the phonology of a language is viewed as important here, so a change in laryngeal specifications merits as much attention as spirantisation, for example, and we will see this abundantly exemplified. We have, in fact, already dealt with a little lenition data here

³³ The questions hardly arise for affricates because they are aspirated in the same way that stops are in English, for example, (see, for example, Kahn 1976, Spencer 1996, Hammond 1999). If laryngeal realism is true for stops, then it is true for affricates as well.

(in section 3.1.3.1.8) and further data presented in section 3.2.1 will widen our understanding substantially.

3.1.4 Prosody

In this section I discuss certain aspects of phonology which are often described as ‘prosodic’. ‘Prosody’ is a complex notion, and different writers understand different things under the term. In what follows, I use the term with a wide frame of reference. The term is intended to cover phonological units and behaviour above the level of the segment and this usage ties in with much current practice (eg, Selkirk 1980, Harris 1994, McCarthy & Prince 1995, McMahon 2000b). This definition shows prosody to be an essential companion to phonological melody, which the last section dealt with at length. The interaction of the two allow for insightful characterisations of certain types of phonological segment, as is shown in section 3.1.5. Despite its fundamental place in phonological models in general, concerns of prosody will not play a crucial role in this thesis. One insight of modern phonological theory is that prosody and melody are ‘different’ and can be considered largely as separate from each other (see, for example Harris et al. 1999, McMahon 2001). This very fact, that the two are different, will play an important role in discussion of aspects of lenition later. Much of the discussion will focus on melodic concerns, in line with the phonological exposition in this section (3.1) Certain fundamental prosodic notions will be vital, however, and one of these will play an important role in the reanalysis of the Germanic processes that were discussed in chapter 2. The discussion here will of necessity be quite brief in places.

To be a little more specific: as well as being the realm of phonological stress and intonation, prosody is concerned with, for example: (i) the status of the syllable in phonology, and hence (ii) the internal structure of syllables and (iii) the positions that segments can take up in syllabic constituents, (iv) the ways in which syllables can be grouped together into higher units such as feet, (v) more simply, the position of segments in words (word-initial, word-final), and also (vi) the quantity or length of segments, to the extent that it can be linked to subsyllabic but suprasegmental timing units.

I assume little that is theoretically controversial here in terms of these points. Syllabic constituents such as ‘onset’, ‘nucleus’ and ‘rhyme’ will be relevant in describing some of the environments for lenition, as will some form of the notion ‘coda.’ Despite the well-known rejection of the syllable in Chomsky & Halle (1968), these concepts are now widely recognised as indispensable in phonology and I assume them here. In common with most

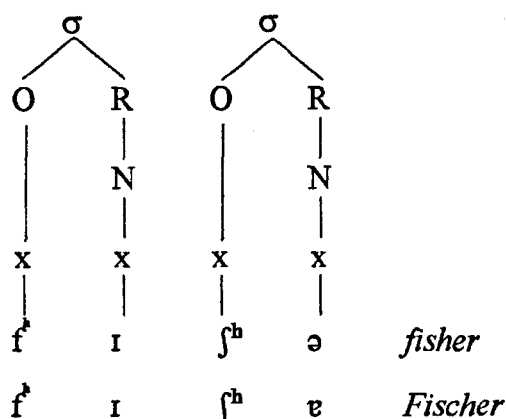
approaches in Metrical Phonology (see for example, Goldsmith 1990), but not standard Government Phonology (eg, Kaye, Lowenstamm & Vergnaud 1990), I notate a syllable directly above the Onset and Rhyme.

I generally show only as much prosodic structure as is necessary in the representations in this thesis. One such potentially vital piece of information is the skeletal timing tier, which forms the interface of prosody with melodic units. I also assume below that syllables are either stressed or unstressed and that this groups syllables into prosodic left-headed ‘feet’ such that lexical stress can play a role in the description and explanation of phonological processes, as can the related notion of the left-headed ‘foot’, which groups syllables together into higher level prosodic units. Finally, I also consider the position of a segment relative to word-boundaries to be prosodic (in the sense that it is non-melodic information), and at times below I refer to simple phonological segmental adjacency.

3.1.4.1 Segments in syllables

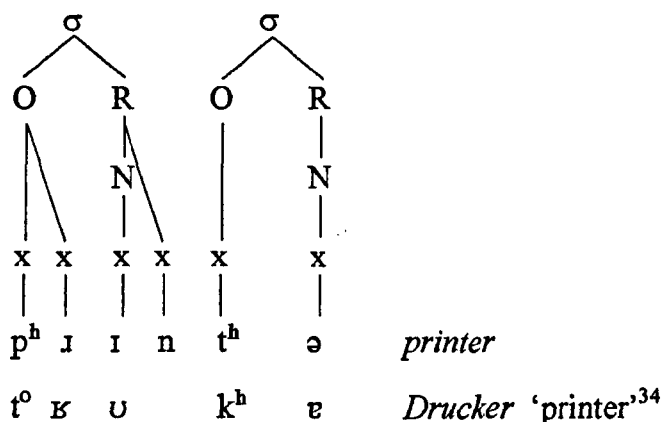
This section contains several exemplificatory syllabic representations of words of Present-Day varieties of English and German, these being well described languages with obvious connections to the historical data from chapter 2. The transcriptions reflect laryngeal clarity, as described in section 3.1.3. The kind of syllabic representations developed here will later principally be used to illustrate syllabic restrictions on the structure of the languages which were investigated in that chapter. The Present-Day varieties are similar in their phonotactics and syllabic structure and will serve as a guide. In one key way, this section explores which prosodic factors condition what is possible in phonology and universalist claims are considered. A word such as *fisher/Fischer* (with English and German pronunciation) can be represented as in (3.19) :

(3.19)



This shows that the word consists of two syllables ('σ') which each consist of an onset and a rhyme. There can be branching in the rhymal constituent as there can in onsets, so a word like *printer* or German *Drucker* 'printer' can be represented as in (3.20):

(3.20)

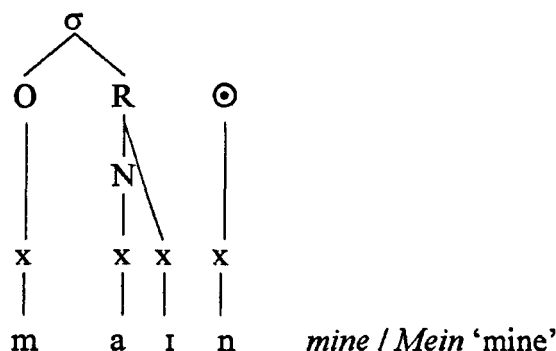


In the German word *Drucker*, only the first onset branches; in the English, both the initial onset and the rhyme branch (giving a coda position in the rhyme, which is not labelled here). The status of final consonants in words is problematic because the segments which can occur in that position have certain special properties (languages often have segmental distributional criteria which only apply to word-final segments, and word-final segments do not always 'count' for the purposes of determining stress) and various proposals have been made to account for this fact (see Piggott 1999 for a recent discussion of the issues). Word final

³⁴ The transcriptions on this page (and elsewhere) look rather different to traditional transcriptions. This is of course because they are largely underlying representations and because they respect laryngeal realism. As mentioned elsewhere, it is not 'normally' necessary to change convention for tasks which simply require the transcription of utterances in [spread] languages - the Roman distinction between the symbols 'b, d, g' and 'p, t, k' can still be used, as long as their phonological worth is implicitly understood. The transcriptions here respect laryngeal realism, however.

consonants are typically either viewed as ‘extrametrical’ in some way (ie, not parsed into any prosodic unit or are parsed into a unit that is higher in the prosodic hierarchy, such as the foot - see Fudge 1969, Goldsmith 1990) or as being parsed into the onset of a final empty-headed syllable (see Kaye, Lowenstamm & Vergnaud 1990, van der Hulst & Ritter 1999a, for example). I take no stance on this issue here, but indicate final consonants, where necessary as as ‘⊙’ (this is a purely ad hoc convention which will not play an important role in this thesis):

(3.21)



3.1.4.2 The timing tier

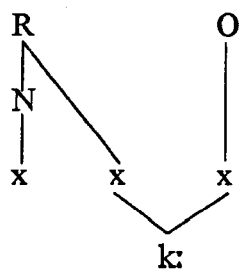
The skeletal timing tier is composed of timing slots, which are marked by ‘x’ in the above diagrams; the notion can be traced back to earlier ‘C’ and ‘V’ slots (originated in McCarthy 1979). Skeletal ‘x’ slots are derived from these ‘C’s and ‘V’s which have been stripped of the slight redundancy of specifying the consonantality or syllabicity of the segments which they dominate (this is already encoded by syllabic position). As is common, consonants and vowels are represented here by the location of the slot under a nucleic or non-nucleic position. The timing tier is part of the interface between prosodic and melodic information and it accounts for the length of segments in a simple, minimalist way. Each segment has a minimum of one slot: short segments are attached to one slot and long segments to two. The use of timing slots competes in phonological theory with the use of moras, which are essentially very like timing slots, but the only occur in rhymes. To a large extent, timing slots are directly translatable as moras (see eg Bickmore 1995 for a comparison),³⁵ although in the ‘mora’ approach, onsets are not allotted any time (or ‘weight’, to use another metaphor - see,

³⁵ At one point in the discussion here I use the symbol ‘μ’, but this is not intended to replace the ‘x’ of the skeletal slot and I explain the reasoning behind it at the appropriate point. Generally ‘x’ slots will be desirable because they can capture the behaviour of geminates which occur partly in onsets, as well as purely rhymal phenomena.

for example, Hyman 1985). This means that one of the insights of the approach to phonological structure adopted here falls onto an unsure footing.

In the case of geminates, the existence of a slot in the onset allows a consistent modelling of their length, in exactly the same way that length is indicated in vowels; vowels are attached to two slots in a nucleus and geminates, which typically occur medially take up the second timing slot in the first syllable and the first of the second, as shown in the pared-down diagram in (3.22). This diagram illustrates a geminate /k/:

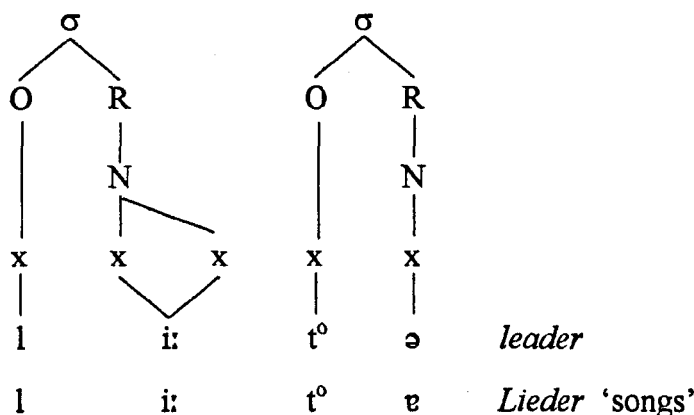
(3.22)



A little more detail will be added to such representations in section 3.1.5, but the basic thrust is that the segment takes timing from both syllables. This rests well with the behaviour of such segments in terms of phonotactics and syllable weight generalisations. They have some properties of single segments, and some properties of clusters and this one-to-two mapping between the prosodic and melodic phonology allows an insightful way of capturing these facts. As we will see below, the basic idea shown here, which is simple nonlinear segmental association, can be extended to account for ‘partial geminates’ which share certain segmental information, but not everything.

The representation in (2.23) shows how this is parallel to the representation of long vowels. The two-syllable words *leader* and *Lieder* ‘songs’ are again highly similar in their phonology:

(2.23)



3.1.4.3 Rhyme Structure: $*\mu\mu\mu$

All of the rhymes so far illustrated have had one thing in common. This is an aspect of prosodic phonology which has frequently been remarked upon and has been thought so fundamental that it has been built into phonological theory on several occasions. This is the fact that none of the rhymes have more than two timing slots. The idea rests on a set of assumptions which are at least partly theoretical in nature (for example, on the formalisation of the 'unusual' status of final consonants as not being in a final rhyme) but these assumptions are not just put forward to account for this observation; they all have independent motivations and would be part of the theory of prosodic phonology anyway, so the generalisation seems to rest on solid ground.

The generalisation partly derives from the observation of static models of the phonology of lexical items, but it also seems to constrain phonological processes, such as compensatory lengthening (see, for example, Wetzels & Sezer 1986 and Bickmore 1995) in that a short segment can spread to take up one additional timing slot if a neighbour is deleted, but no more than one. Syllables with more than two skeletal slots in the rhymes have at times been argued to exist in certain languages (they are called 'superheavy') but where they do exist, they are often unstable and disappear diachronically, or require special licensing mechanisms.

The insight concerning 'only two slots in the rhyme' has been expressed (i) in early moraic work (eg, Hayes 1986), (ii) in work in Government Phonology, where it is formalised as a basic of the theory (eg Kaye, Lowenstamm & Vergnaud 1990), and has been taken over into developments of this model (eg, van der Hulst & Ritter 1999), (iii) in Optimality Theory, where it has been formalised as $*\mu\mu\mu$, which means that it is forbidden for a syllable to have

three or more moras,³⁶ (see Prince & Smolensky 1993 and also Bermúdez-Otero 1999), (iv) by Vennemann (1988) as a preference law for syllable structure: “in stress accent languages an accent syllable is more preferred the closer its syllable weight is to two moras” (1988, 30) and (v) there is also a tradition of recognising a principle of this kind in philology and historical phonology (for example, Lass 1994, 36-38 reports how philologists have described geminations as only able to occur where they do not create superheavy syllables).

The basic principle is well established and it seems that this, if anything, is a naturalness principle which we might expect to guide the innovation of phonological processes. In what follows, I call it $*\mu\mu\mu$, taking over the formulation of Prince & Smolensky. This terminological choice is certainly not intended to imply historical precedence of the OT formulation, nor any commitment to the derivational machinery in the framework. It is simply a compact way of formulating the principle. It is the only way in which I use the mora symbol ‘ μ ’ rather than the skeletal-slot symbol ‘ x ’, but this is simply because, for me, ‘ μ ’ means ‘ x -slot in a rhyme’. It is more compact than writing ‘the principle that there cannot be more than two x -slots in a rhyme’. In what follows, I show how a consistent recognition of $*\mu\mu\mu$ can help us to understand aspects of the historical phonology of German.

3.1.4.4 Summary and prospect

The discussion of prosodic phonology in this section has, at times been simplistic. However, the issues that I have simplified here would not substantially change the conclusions that I reach below if they had been investigated in greater detail. As elsewhere in this thesis, there is a need to set aside certain issues where the points involved do not contradict or disprove the position I adopt. Nonetheless, the set of assumptions that have been outlined here for prosody will form an important part of the general picture of phonology that I come to apply to the question of what lenition really is, and hence, what the GCS, HGCS, EIFV and IGCW really were.

In the next section, the syllable structure trees and skeletal timing slots that were exemplified here will be shown to intersect with the elements of segmental structure and to provide a situating point which groups them together into lexical units. The bonds that can be seen to be formed over this structure will prove to be a vital aspect of our understanding of lenition.

³⁶ Given the fundamentals of the theory, this constraint is violable in OT, but the fact that it has been formulated at all indicates that theoreticians want to be able to use it.

3.1.5 Elements in segments

Section 3.1.2 argued that segments play an important role in phonology. Section 3.1.3 argued that these segments are composed of a specific set of privative elements. This section shows how the two notions combine by demonstrating the elemental composition of several key segments which featured in chapter 2 and which will feature again in chapter 4, as well as certain segments which are discussed in section 3.2. Some of the notions considered in section 3.1.4 will also be considered here.

The simple definition of the phonological notion ‘segment’ adopted here, in common with other work, is ‘one root node and all elements that are attached to it’. The segmental representations presented in this section will require a few basic structural assumptions. The first of these concerns the organisation of elements within segments. Firstly, in common with much other work, I assume that the subsegmental units are attached to a root node. In addition to the assumption of a root node, it is conventional in much theoretical phonology since Clements (1985) and Sagey (1986), to assume that elements, or their equivalent, are organised under a number of class nodes in a geometric tree which are directly linked to the root node. This is not so conventional in element-based approaches, although Ewen (1995) explains how dependency relations have a similar effect in Dependency Phonology and Harris (1994) and Brockhaus (1995) assume a simple element geometry. For the purposes of this thesis, geometric representations would not make any different predictions than would segmental representations with unordered sets of elements (as in Kaye, Lowenstamm & Vergnaud 1989 and Harris 1990). I therefore generally use such simple representations with elements linked to one basic root node in any order, although simple geometric representations could reorganise the elements under class nodes and not affect anything; such representations can be thought to underlie the representations used here. One exception to this generalisation will become apparent later when it will be seen that segmental representations can encode the relative timing of conflicting elements.

The second point is connected with another common assumption in theoretical phonology. This has to do with the potential for different kinds of association between the melody of segments and the timing slots of prosody. We saw in section 3.1.4.2, for example, that the existence of the root tier and timing tier allow us to characterise geminates in an insightful way, and this kind of representation will take on a real importance below.

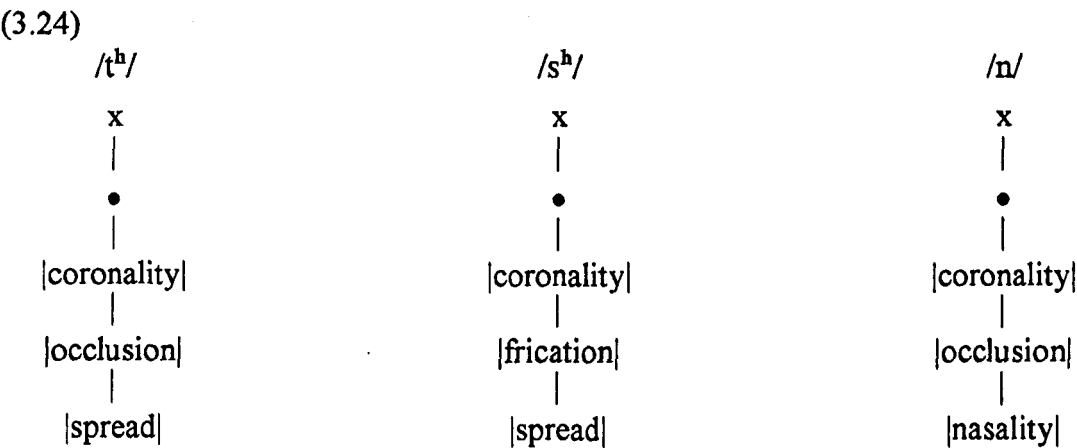
As explained in previous sections, elements are assumed to behave in a ‘nonlinear’ fashion in that assimilations are accounted for by elements spreading from one segment to others; this means that one element can be attached to more than one segment. This principle also allows

for the lexical sharing of elements by more than one underlying segment (that is, sharing which is not derived by a process in the phonology) which can account for phonotactic co-occurrence restrictions and other aspects of phonology. Given the definition of ‘segment’ adopted here, such an arrangement would still count as two underlying segments, however, as there would be (at least) two root nodes.

Exemplification of segments

The segmental representations provided in this section are given for segments composed of the set of six manner and place elements and three laryngeal elements presented above. Representations are only given for obstruents, apart from one nasal segment, which is included for the sake of completeness, and for approximants, as these will feature in one important aspect of the discussion below. The segmental symbols are used in accordance with the proposals for laryngeal realism made in section 3.1.3.1 which take account of the segments’ laryngeal specifications. Headedness is represented only where more than one segment has the same set of elements.

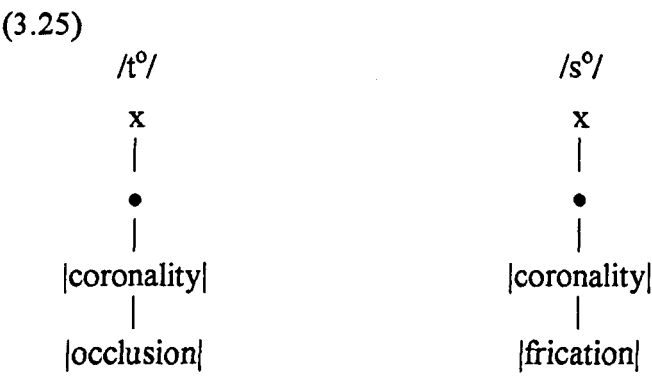
The representations in (3.24) show how three basic segments are characterised in |spread| varieties of English and German:



The representations include an ‘x’ for a timing slot and a ‘•’ for a root node. Unless a certain geometry were to be assumed, each of the elements could be considered to be attached to the root node individually, although this is not made explicit in the diagrams in (3.24). The segments in (3.24) are typically described as the ‘alveolar voiceless stop’, ‘alveolar voiceless fricative’ and ‘alveolar nasal’, although, as we saw in section 3.1.3.1, these descriptions can be misleading. The obstruents are indicated with |spread|, as there is good evidence that most

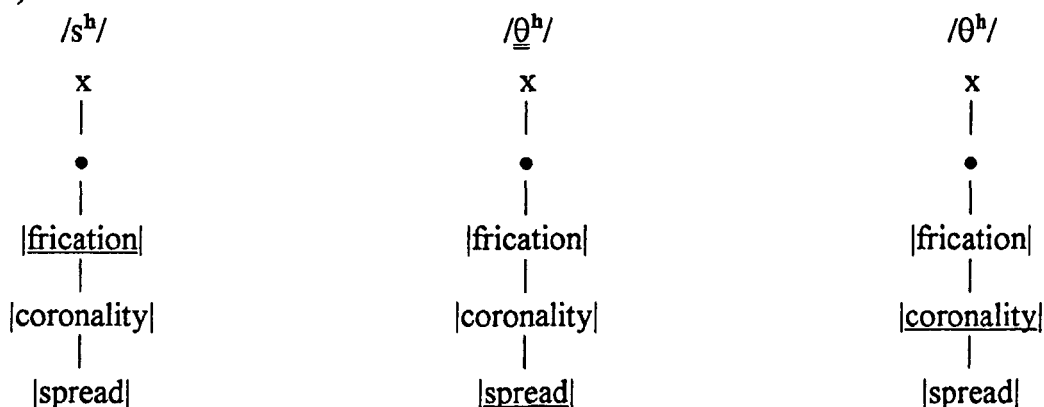
varieties of English are |spread| languages, as we saw above. Segments such as /d/ and /z/ have exactly the same structure as the first two segments here, except they feature |voice| in the place of |spread|.

The representations in (3.25) are for the ‘neutral’ stops and fricatives that occur in both |spread| languages and |voice| languages; these would share place of articulation with those in (3.25) but would contrast with them in terms of (their lack of) laryngeal specifications:



The role that headedness can play in a segment is illustrated in (3.26). This shows representations of the segments /s^h/ and /θ^h/, which contrast in Present-Day English, and a representation for the segment /θ̥^h/ which was described in section 2.3.1. These all feature the same elements, but are perceptibly different and hence can contrast. The difference is accounted for by headedness, in that a different element is the head of each of the segments. The heads of the segments are underlined, following the convention of Harris (1994):

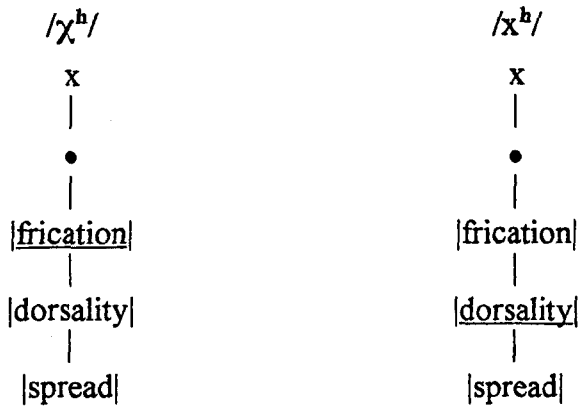
(3.26)



The decision as to which element is the head of which segment is not arbitrary, but is based on observable phonetic behaviour. The element |frication| is the head of /s^h/ because the segment shows greater stridency (ie, noisiness) than the other two; the noise of frication is more pronounced in this segment than in the other two (this reasoning is akin to that of Harris 1994 although he is working with a slightly different set of elements). The fact that |spread| is the head of /θ/ ties in with the fact that the oral articulation of this segment is less extreme than for /s/ or /θ/, so we would expect that a non-oral element would be the segment's head. /θ/ is a “controlled articulation” (Hickey 1984, 234) but there is less contact with the roof of the mouth than for /s/ (as Pandeli, Eska, Ball & Rahilly 1997 show through electropalatography) because it has a broad central channel and does not involve tongue grooving (in contrast to /s/); it also does not feature forward movement of the tongue towards the teeth (in contrast to /θ/). It is no surprise that three fricatives can be made at this place of articulation and that three elements are involved, which leaves the third, |coronality| as the head of /θ/.

The same kind of distinction (on the basis of headedness) can be made for the slight difference between labiodental and labial fricatives, and between velar and uvular fricatives. The latter distinction is shown in (3.27):

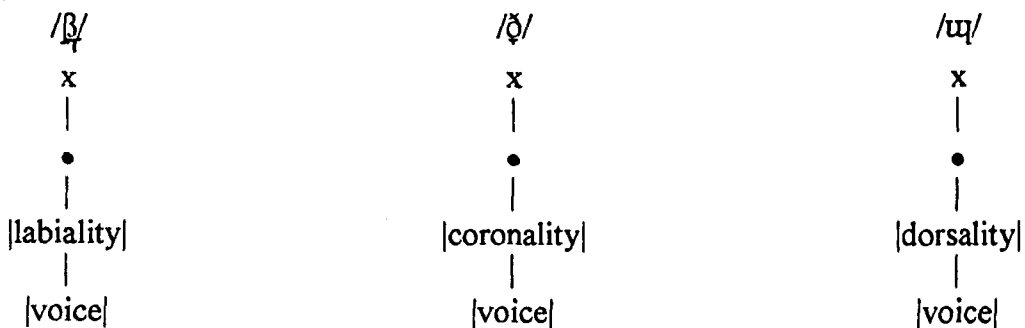
(3.27)



Similar reasoning underlies the choice of heads here: the uvular is the more noisy of the two. It is possible here that physiological constraints prevent the formation of three different fricative here (ie, it is not possible to form a clear groove in the tongue in the dorsal area, whereas it is possible at the alveolar area).

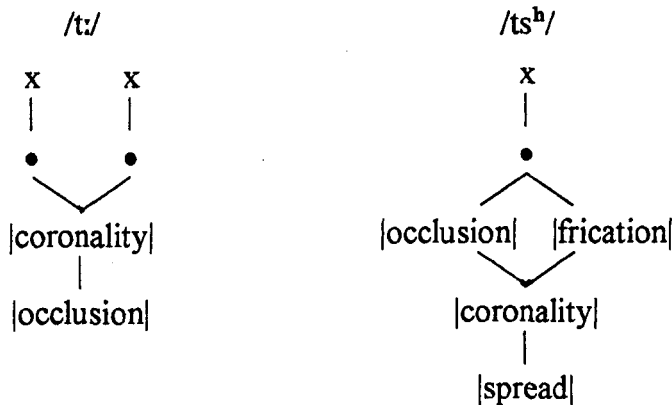
While they are not strictly obstruents, we can note that the elemental approach which we have adopted here also allows for the simple characterisation of approximants. Approximants involve a definite movement of an articulator, but no substantial stricture, so they can be simply represented as the ‘place’ elements alone, without a ‘manner’ element. The representations in (3.28) show three approximants with a |voice| specification:

(3.28)



The next set of representations exemplify some slightly more complex segmental representations. The representation of the geminate is as in section 3.1.4, illustrating the interaction between melodic elements, root nodes and timing slots. We saw in section 3.1.4 that, among other things, timing slots encode the length of segments. The example is a laryngeally neutral geminate. The second segment in (3.29) is an affricate, and is represented in what looks like a similar way to a geminate, but is actually quite different.

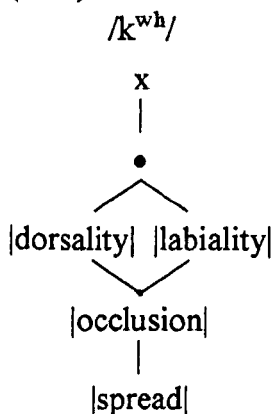
(3.29)



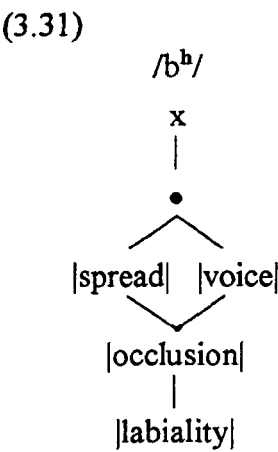
How does the representation for the affricate fit with the claim that was made in the introduction to this section that elements can generally be viewed as an unordered group? Clearly the notation here indicates that the occluded part of the affricate precedes the fricated part when the segment is produced and that seems to be the right result, but if the elements are truly unordered, then that effect is just a trick of the formalism. It seems here that there must be slightly more than entirely minimal structure to the segment, such that the relative timing of the elements can be encoded. This is not unreasonable, given that the two elements |occlusion| and |frication| are mutually exclusive because they make incompatible claims on the articulators. In just this kind of case, then, segments such as affricates, which are often called ‘complex’ or ‘contour’ segments, clearly have the capacity to encode the information which is necessary to allow the segment to be pronounced, as is quite commonly assumed (see, for example Sagey 1986 and the discussion in Lombardi 1990).

The IE labiovelars, on the other hand, do not represent such a problem. They can be represented as in (3.30), which makes further use of the ‘complex segment’ formalism, but in this case the two elements are compatible with each other and can both have an effect at the same time. The segment is pronounced with concurrent velar articulation and labialisation and the formalism can be seen as just that, a formalism intended to bring the double articulation to our attention:

(3.30)



In section 3.1.3.1, we saw that segments can feature more than one laryngeal element. This, too, is easily represented, as shown in (3.31), where the two laryngeal elements are interpreted in the same way as the two place elements in the labiovelar in that they do not need to be seen as chronologically ordered:



The segmental representations proposed here are partly standard and partly somewhat novel. The novelty will become further apparent in the coming analyses which make use of these representations. The kinds of representation given here are all that we will require for the coming discussion of lenition.

3.1.6 Summary and prospect

The model of phonology presented in this section (3.1) adopts a number of ideas that have become standard in a range of phonological models. The references cited when these principles were first introduced and subsequently discussed illustrate the range of influences that have shaped the model presented here. As was explained in section 3.1.1, the model is based on a metatheoretical position which predicts that phonology will be minimalist, to the extent that over-abstract analyses and analytical machinery is deemed unlikely. This basic stance has similarities with those of others, of course.

It is clear that this constraint of minimalism plays a similar role to the idea of ‘naturalness’ which was explored in section 1.2.2.2, but in fact, the two ideas do not quite place the same constraints on phonology. It is expected that notions of naturalness constrain the novel innovation of a phonological process, whereas the minimalist idea means that, even once a process has been newly innovated into the phonology of a language, it cannot change to such

an extent that the underlying segments become very different from their surface counterparts. It is this that can lead to a relatively 'rapid' lexicalisation of a phonological process.³⁷

The model of phonology expounded here raises and addresses questions which are interesting in their own right, but in the rest of this thesis, it will be used as a tool to help us understand obstruent lenition processes in general and the data from chapter 2 in particular. We will return to discuss certain phonological issues in section 3.2.2, where we see how others have applied phonological theory to the study of lenition and subsequently investigate how the model adopted here allows lenition to be understood.

The details of this model, explored with an aim to be minimalist, were set out for melody in section 3.1.3 and, much more briefly, for prosody in section 3.1.4. Aspects of them were brought together in section 3.1.5. One underlying concern in this section has been the place of the segment in phonology, and this concern led to the proposition of a clarified set of phonological symbols (or, rather, symbol and diacritic combinations) which will allow us to be as clear as we can about the reality of the segment under discussion. I use these in the coming discussion of the patterning of lenition processes.³⁸ These depend on the recognition of the set of laryngeal elements which was proposed in section 3.1.3.1. These elements were robustly defended there, and this involved considering both the general possibilities in language in terms of laryngeal phonology and their specific application to certain languages which are highly relevant to this thesis. The rest of section 3.1.3 dealt with place and manner and was more universalist, largely simply showing what the possibilities are.

When we begin to reconsider the processes which were investigated in such philological detail in chapter 2, we will not find the effort put in there to have been wasted. While the phonological issues raised in this chapter are vital to our understanding of historical phonology, they are worth nothing without a firm philological background. As we saw in the discussion of the 'fortis/lenis' distinction in section 3.1.3.1.4, the observational skills brought to language history in philology are keen, and essentially recognised many of the points behind the laryngeal realist position. The final analysis of the processes discussed in chapter 2

³⁷ The notion of 'rapid' here is meant to make a comparison with what is allowed in certain other models of theoretical historical phonology, such as relatively 'standard' generative ideas, where highly abstract underlying representations were allowed (see Bynon 1977, Kiparsky 1988, McMahon 1994). It is not meant to imply that phonological processes are always quickly lexicalised. Certain highly motivated 'transparent' phonological processes can survive for centuries, of course.

³⁸ As was mentioned above, the assumptions made here would, strictly speaking, have a substantial impact on transcription, thus *dig* is /t^oik^o/ and *pub* is /p^hup^o/ (with Northern British English vocalism). For the purposes of pure transcription in any individual language, however, there is no need to respect laryngeal realism, and it makes sense to continue as before. The laryngeal realist symbols become vital when universalist or cross-linguistic phonological claims are being discussed.

will be quite different to the initial presentation, however, as it will be informed by the phonological ideas proposed here, and also by a detailed study of lenition, which is to come directly, in section 3.2. The kind of processes involved in the data from chapter 2 will be subjected there to considerable investigation. We will also consider certain aspects of the phonological environments where lenition processes occur.

We have yet to consider in detail precisely what phonological processes do to the phonological representations that were developed in this section, or the notions that are tied up with the innovation of such processes. In section 3.1.1, I explained that elements can be lost from the make-up of a segment, and we saw in section 3.1.3.1.8 how this allows for an insightful characterisation of debuccalisation. We return to both debuccalisation and the idea of element loss in section 3.2.2. Strictly phonological processes can also spread elements from neighbouring segments. These are the phonological possibilities which can motivate, or allow for, the innovation of phonological processes by one quantum (but perhaps not strictly explain them, as we saw in section 1.3.1). As was also briefly discussed in section 3.1.1, and in somewhat more detail in section 1.2.2.2, strictly phonetic factors can lead to the innovation of processes, too. This means that both ‘phonological’ and ‘phonetic’ factors can provide the necessary potential impetus which allows for novel phonological processes to be endogenously innovated. They only allow for certain types of processes, however. I return to this point in section 3.2.2.3, where we will see that this allows for certain predictions to be made about what is a possible process, and perhaps more importantly (given the discussion of the notion of explanation in section 1.3.1) about what is not a possible process, so that we might be able to explain and understand the types of process that we have described here as obstruent lenition.

3.1.6.1 Relationships between processes

In the model of phonology discussed in this section and in the general model of phonological change adopted in this thesis, there is not really any place for a psychologically real relationship between separately innovated phonological processes. This goes against the idea which has been proposed by some historical phonologists that individual processes (such as those provided for by the phonology discussed in this section) can be related to each other in a perhaps chronologically explanatory sense.

Martinet (1955) influentially claimed that processes can be linked in ‘push chains’ and ‘drag chains’. This idea is founded on the proposal that the occurrence of one process can in some sense be seen to cause another, so, for example, if we can see that there has been both a

spirantisation in a language of the type /b, d, g/ > /β, ð, γ/ and a voicing of the types /p^o, t^o, k^o/ > /b, d, g/, it might be proposed that spirantisation came first and ‘pulled’ the voicing after it, to replace the segments that were lost in the spirantisation. Or it might be claimed that the voicing began to occur first, but that it was realised that this would involve a merger of six segments to three, so instead this process ‘pushed’ the existing voiced stops to spirantise. Indeed, stories of this kind have been proposed for the set of changes involved in the GCS (and also for some Spanish data which we will discuss in section 3.2.1.2).

I do not consider such putative relationships in this thesis. From the above, it will be clear that it is generally not possible to tell, given the tenets of ‘chain theory’, whether a particular chain in the history of a particular language should be viewed as a push chain or a drag chain. In terms of process innovation it then becomes impossible to talk of explanation in any sense, partly because we can’t tell which way around the causation is supposed to have occurred, and partly because the notion of causation used here is directed towards the influence that underlying segments have on each other in a phonological system. This rejects any place for the kind of phonological and phonetic natural and minimalist reasoning for process innovation which was developed in section 1.2.2.2 and in this section. On the view developed here, then, chain theory is rejected as a causal notion. It is perfectly possible for a linguist to use chain reasoning to see a connection between a set of processes, but languages and their speakers do not.³⁹

The way in which the understanding and explanation of historical processes is approached in this thesis is to focus on the phonetic and phonological motivation for them. It seems to me that if we are satisfied with an ‘explanation’ such as, for example, d > ð *because* previously (or even synchronously) t > d, then we will in fact fail to start looking for the real reason. In order to avoid this, I focus on the segments and the processes themselves.

3.1.6.2 Elements and segments: particles and atoms

In the coming discussion of lenition, the segment types recognised in this section (3.1) will form an important part of the discussion. We will see that certain types of segment behave diachronically in certain ways and others in different ways. While this is not always

³⁹ This could be seen to fly in the face of my grouping processes together in chapter 2 under headings such as ‘the GCS’ and the ‘HGCS’, and indeed it does. As was mentioned at the start of section 2.1, the groupings adopted in this thesis are conventional ones, motivated by the traditions of the philological literature. As we will see below, some of the processes involved are very different from each other.

necessarily a vital correlation of the recognition of the different synchronic segment types, it will provide a reassuring corroboration.

Some of the ideas which were discussed in section 3.1.1 imply that the elements which were introduced here should not be thought of as part of a specifically phonological UG. This seems the right decision. While the elements are cognitive units, as part of phonology, they are glossed in terms of the articulatory impact that they have and it seems entirely possible that they could be gained in acquisition through induction over the inherent possibilities in the vocal tract. We need to assume that the bundles of muscles, and the possible contrastive actions for those muscles in the vocal tract only allow us to abstract certain elements from that which we acquire, and this seems plausible. There are only certain things that we can do. The status of entities such as * $\mu\mu\mu$ is not so clear. While its place in the phonology of the world's languages is assured, as shown in section 3.1.4.3, we would need to postulate a physiological or general cognitive constraint from which it could be derived if it is not directly provided by the genome. I leave this question open here.

It may be interesting, for a moment, to consider the status of elements in regard to segments. We have seen that at least certain elements have stand-alone interpretability (although this has not been demonstrated for them all). Most segments, however, are composed of more than one element. The use of the term 'element' to describe these subsegmental particles was taken, as explained during the discussion, from the tradition of Government Phonology. The name is an interesting one and could be seen to make metaphorical claims which extend from the field of chemistry, where the term seems to be taken from. In fact, it is not clear to me that 'element' makes for the best metaphor. Durand & Katamba (1995) talk of 'phonological atoms' and this is perhaps a step in the right direction metaphorically. In fact, chemistry offers us a better metaphor. The phonological elements discussed here seem best compared to subatomic particles such as the proton, neutron and electron. These combine to make up a range of atoms, which we might equate with the segment (note that a single stand alone proton is interpretable - as hydrogen). Of course, atoms combine to make up molecules, so we might hope that there will be a phonological equivalent of the molecule to make the metaphor complete (although we probably should not let ourselves be blindly led by metaphors).⁴⁰ We will see in section 3.2.3 that there is indeed such a type of entity in phonology. To discover what that is, we must turn to the specifics of lenition processes, and this is the subject of the imminent second major section in this chapter.

⁴⁰ I return to the role that a metaphor can play in phonology later in this thesis.

3.2 Another look at lenition

As this thesis has progressed thus far, I have developed a set of ideas which, while connected in principle, have not yet clearly been connected in practice. In section 1.4, a simple idea of lenition was introduced which dealt with types of segments (exemplified by a set of standard segmental symbols) and the relationships that hold between them in lenition processes. In chapter 2, certain attested processes from the history of Germanic languages were introduced, with a gradually increasing level of phonological sophistication. In section 3.1, a model of phonological analysis was introduced which partly involved the description of a prosodic means of organising segments with relation to one another, but also, and more importantly, a model of subsegmental structure was introduced in terms of phonological elements, some of which are very similar to individual features in feature geometric models and some of which are not so clearly similar.⁴¹ The recognition of the set of elements and the patterns of their usage in various languages in that section led to the awareness that, where comparisons are to be made in terms of the behaviour of segments across languages (ie, to the extent that we make universalist claims), we must compare like with like. For this purpose, a set of unambiguous symbols was introduced which respect laryngeal realism.

Certain aspects of the points made in three above-mentioned sections conflict with each other. As was briefly touched upon at the end of section 3.1.3.1, the recognition of laryngeal realism means that it is at least possible that such formulae as ' $k \Rightarrow x$ ' have become problematised. We may wonder, for example, whether this process can be innovated when the ' k ' (that is, the T) is a $/k^h/$ and when it is a $/k^0/$. Is $k^h \Rightarrow kx^h$ a natural process? Is it possible through natural endogenous linguistic means to innovate $k^0 \Rightarrow kx^h$? In terms of these processes, we may well also wonder whether it is possible to innovate ' $k \Rightarrow x$ ' in one step, in any case, in line with the notion of the possible quanta of innovation, as discussed in section 1.2.2.3.

These questions and more of the same general type are addressed in this section (3.2). The whole focus of the section, which is really the second half-chapter of chapter 3, is on lenition. Much of the discussion, especially in section 3.2.2, is on lenition processes in the abstract; the focus there is on general process types and it is in that section that my basic understanding of

⁴¹ In and of itself, the fact that the similarity is only partial does not speak for the superiority of either model, but I have argued on minimalist grounds that it is better to have units which all have fundamentally the same nature than to have a set of units which are fundamentally heterogeneous.

precisely what lenition is (as a general notion which groups together different kinds of phonological processes) will become apparent.

There is an empirical aspect to this section as well, however. In section 3.2.1, after I explain why there is a need to do this, I briefly investigate a range of other attested lenition processes. The treatment here is much less detailed than it was for the Germanic processes in chapter 2 and full analyses are not proposed for all the data mentioned, but the discussion in this section will allow an important widening of the empirical base to the thesis. The investigation of the further examples of lenition processes will partly help to inform the generalised discussion of lenition processes which follows it, but it will also help with the treatment of the interaction that can be observed between lenition processes and phonological environments. As was noted in section 1.4 the two notions of phonological process and phonological environment can be separated from each other, but they do frequently interact, of course. The treatment of environmental effects which is adopted here (in section 3.2.3) is quite different to most previous treatments because it does not focus on the idea that certain phonological environments ‘promote’ lenition; rather, the focus is placed on the idea that some environments can be seen to inhibit it. The more abstract discussion in this half chapter is followed in chapter 4 by the application of the notions developed here to the specific cases of lenition for Germanic languages that were first introduced in chapter 2.

The focus in this section is thus on ‘lenition theory’.⁴² We return to the lenition trajectories given in section 1.4 and repeated here as (3.32)⁴³ in an attempt to understand what they really mean.

(3.32)

$$k \Rightarrow kx \Rightarrow x \Rightarrow h \Rightarrow \emptyset$$

$$k \Rightarrow g \Rightarrow y \Rightarrow \emptyset$$

Given the notion of the phonological process which has been developed in chapter 1 and section 3.1.1, we might expect each of these steps to involve a phonological process. It is claimed (eg in Lass & Anderson 1975) that these processes are observed again and again in a wide range of languages, so each step down the trajectory should be formalisable by itself. In

⁴² I have taken this term from Hickey (1984); while it does not seem to be in wide use, it is an apt one to describe the general understanding that a theorist has of how lenition processes should be analysed and of what they have in common that allows them to be grouped under the name.

⁴³ Because they are repeated from 1.4, the symbols in (3.32) do not respect laryngeal realism, but this does not affect the points made here, and the trajectory will be reformulated in due course so that they do.

this thesis it has been argued that processes can be innovated either through simple ‘natural’ articulatory phonetic or phonological innovation or through a reanalysis of acoustic aspects of the signal. In this way, phonological processes are expected to be formalisable as involving either element loss or gain (from a local source) or, if they involve any other process, this must result from a simple reanalysis of acoustic or articulatory pressures. We can anticipate that it is these kinds of factors that will be the ‘cause’ of the lenitions⁴⁴ and that they should be the basis of lenition theory.

Some of the discussion in the first half of this chapter (section 3.1) may seem to completely call into question the concept of lenition (as it was presented in the preliminary section 1.4) as something that can be said to truly exist phonologically, in that it is not immediately obvious how an overarching concept such as ‘lenition’ can fit in with the minimalist ideas developed there. By the end of this section we will see that lenition as a concept is not a completely bankrupt notion, but it will only survive in a radically reinterpreted form. Still before the interpretation of the term is clear, however, we turn to some further instances of attested processes which would generally be referred to as cases of lenition.

3.2.1 Lenition in other languages

In this section I broaden the empirical base for the investigation of the properties of lenition processes. This is a modest increase of scope, largely for reasons of space. Given the fact that one key aim of this thesis is the detailed study of four sets of Germanic data, it would not be helpful to engage in detailed discourse on processes from other languages; however, it would be unwise to generalise both (i) in the abstract, and (ii) with specific reference to the Germanic data, on the basis of only those processes which are already under investigation. Some of the lenitions described in this section are extremely well-known (rather like some of the data from chapter 2) and this has in fact been one of the principles guiding the selection of the data here. The section thus includes some data from Spanish and Celtic which are frequently brought into the discussion of the concept of lenition. Another reason behind the selection of the data discussed here is a goal that the data should be relatively diverse, both in terms of the genetic affiliation of the languages involved and in terms of the processes that they illustrate. They are

⁴⁴ The word ‘cause’ is in scare-quotes here because of the discussion of the notion of cause, prediction and explanation in section 1.3.1. It might be better to write that such factors are the ‘key contributory endogenous factors’.

partly chosen to complement the processes found in the data in chapter 2 by exemplifying some common types of lenition not featured there, and they are partly intended to supplement the previous data with analogues. As was mentioned in section 2.3.1, an important analogue for the HGCS₁ is included in this section, from Liverpool English.

The data from Liverpool English are the only original data presented here, the rest, as was largely the case in earlier sections, are gathered from either philological sources or, at times from more ‘phonological’ work. This partly reflects the fact that the processes covered in this section include clearly synchronic processes as well as diachronically fossilised processes (that is, processes which can be reconstructed from the effects that they had on the phonology of a language once they were lexicalised). This illustrates the interaction between synchrony and diachrony in lenition processes. The processes described in chapter 2 were once synchronic processes but are now entirely fossilised and it may well be instructive to examine some currently synchronically investigable lenitions

The discussion of the processes in each of the languages will be quite compact, but it is intended to be nonetheless detailed enough to be useful in the discussion that is to follow. This requires a certain amount of discussion for each of the cases involved and it means that a simple one or two line description (such as would be all that is possible in a broad-sweep approach) is not suitable here. Lavoie (1996, 2001) and, following Lavoie, Kirchner (1999) include extensive lenition surveys in their treatments of the concept. These are admirably broadly based surveys and are useful for considering general patterns, but due to the very nature of such surveys, the treatment of each individual case is quite superficial. If key details for the environment of each process are to be considered, along with the implications of laryngeal realism and the question of what implications there are if a processes is no longer synchronically active, it is more helpful to deal with a smaller number of processes in greater detail.

In this section, for the above reasons, I deal to differing extents with data from Spanish, Proto-Iranian, Dravidian, Celtic, Greek, Bantu and Liverpool English. This discussion forms the empirical basis for the broadening of the discussion of lenition processes. Later in the thesis, I also touch on aspects of the phonology of varieties of Italian and Dutch, where these are shown to fit with the patterns established here.

To the extent that it is possible in a brief discussion, I consider in this section whether the obstruents in the cases which are to be addressed would have featured [voice] or [spread]. I also discuss aspects of both the phonological process involved and of the environment which

conditioned it. These latter two points will feed into the discussion in sections 3.2.2 and 3.2.3, respectively. In the discussion of the processes, I chiefly make use of the whole segment symbols which were introduced in section 3.1.3.1.7. As explained in section 3.1, these are understood to be comprised of a set of elements but also to have an existence in their own right. Just before we engage with the data, I present a brief consideration of the notion of phonological environment and of those environments which are likely to be of interest here.

3.2.1.1 A brief excursus on environments

As noted in section 1.4, a connection between lenition and ‘lenition sites’ has been recognised by certain researchers; in fact there is some mention of a conditioning environment in most writing on the topic, although the importance that different writers accord to the connection between process and environment varies greatly. One recognition in this thesis has been the point that the phonological processes involved and their environments are separable but it would be overhasty to deny that there is any connection. We saw in the discussion of the GCS₁ and HGCS₁ in chapter 2 that environment can indeed play a role in lenition. As was recognised in that section, there are certain ‘exceptions’ to the processes; these illustrate certain phonological environments which affected the introduction of the lenitions.

Several generalisations emerge from the discussion of environmental factors in previous work on lenition where an emphasis is placed on the issue (for example, Escure 1975, 1977, Harris 1990, 1994, Kirchner 1998, Ségéral & Scheer 1999). In section 1.4, it was mentioned that relevant ‘lenition sites’ have been claimed to be ‘medial and final positions’ but these are vague notions. Although the distinction is not always recognised, these generalisations can be seen to be connected with both prosodic and melodic concerns. In terms of a segment’s prosodic environment, we can recognise: (i) the relationship of a segment to syllable boundaries, (ii) the relationship of a segment to word boundaries, (iii) the relationship of a segment to stressed vowels. In terms of melodic environment we can recognise: (i) the nature of the preceding segment (in terms of place, manner and laryngeal elements) and (ii) the melodic nature of a following segment.

Bearing these concerns in mind, we can abstract several phonological environments which may prove helpful in the discussion of lenition and environment; these are shown in (3.33),

along with a simple gloss and a key for the symbols used.⁴⁵ Some of the environments are given two glosses; this either illustrates the fact that they could represent an environment where either prosodic or melodic concerns might be relevant or that there is more than one terminological tradition:

(3.33)

- | | |
|---|------------------------|
| A [__#] - 'word-final' | |
| B [__c] - 'coda' or 'pre-consonantal' | # = word boundary |
| C [v__v] - 'intervocalic', 'medial' | c = any consonant |
| C1 [ˇ__ (v)] - 'foot-internal', 'post-stress' | v = any vowel |
| C2 [(v)__ ˇ] - 'foot-initial', 'pre-stress' | ˇ = any stressed vowel |
| D [c__] - 'onset' or 'post-consonantal' | |
| E [#__] - 'word-initial' | |

Discussion of the implications of the interaction between lenition and its environment will be taken up later in this thesis (in sections 3.2.2 and 3.2.3) and I will not pre-empt this here. However, we can note at this juncture that the nature of the interaction is not pre-theoretically obvious; that is, we can see that certain processes have been introduced into the linguistic system under investigation in certain environments, but we do not know whether it is the case that (i) certain environments have 'promoted' the process or (ii) the complementary set of environments have 'inhibited' it.

Whichever way of viewing the effects is the right one, however, some previous work on lenition might lead us to expect lenition to occur in environments A, B and C, but not in D and E. Some work proposes that lenition principally occurs 'intervocalically' (for example Trask 1996: "above all between vowels") and here only environment C would be expected to have an effect. Or if lenition is thought to occur 'postvocally' (a term used in Labov 2001), this could include A, B and C.

Environment C has often been split (for example in Harris 1990, 1994) to take account of stress placement and foot structure, and this is offered as an option in (3.33). Where this is

⁴⁵ It is perfectly true that some of these could overlap (for example D and C2); it is important in such cases to recognise the duality of the environment; it is also true that the environments given here are not entirely exhaustive in that we might well imagine other phonological environments which could possibly have an effect (no consideration is given to 'long-distance' environments, such as are relevant in vowel harmony, for example), but they will serve us well for discussion and are taken from a distillation of the lenition literature; the inspiration comes especially from Ségéral & Scheer (1999).

done, it is typically claimed that lenition might be expected in C1, but not in C2. For example, ‘tapping’ is a synchronic process which occurs in many North American, Australian and Irish varieties of English. It sees an underlying alveolar stop surface as an alveolar tap (thus: /t/ → [ɾ]) and is frequently seen as a type of lenition (see Harris 1990, 1994, Trask 1996) because of the reduction in articulatory force that it required. While I do not deal with tapping further (because the output is a rhotic, not an obstruent) it is mentioned here as a classic case of a process which is sensitive to foot structure: it occurs in C1 but not in C2 (in *pretty* [ˈpɹɪɹi] and *water* [ˈwɑɹə] but not in *pretend* [prəˈtend] and *boutique* [buˈti:k] - data from Harris 1994).

So, all the environments in (3.33) have been singled out for attention in previous treatments of lenition and it is thus possible that they will be of use in the coming discussion of further examples of lenitions in this section. Accordingly, the treatment here will include a consideration of these environments where it is appropriate (some languages do not allow obstruents to occur in all these positions, so naturally certain environments are not relevant in all cases). Some of these environments were investigated in the presentation of the data in chapter 2 and they will feature again in the reinterpretation of that data in chapter 4. As will be apparent from the discussion that occurred there, however, and further discussion to come, some of the processes in chapter 2 were only affected by phonological environment to a limited degree.

3.2.1.2 Spanish

Some of the best known lenitions have occurred in Romance languages (in particular, the western Romance languages, see Posner 1996).⁴⁶ As an example of ‘Romance’ lenitions, I describe in this section at least part of the situation in many varieties of Present-Day Spanish. This will involve a brief discussion of both historical and contemporary lenitions in and of themselves, but the main focus, as elsewhere in this supersection, is on what an investigation of

⁴⁶ In fact, it is quite possible, as Posner (1996) discusses, that certain common lenitions in what are now individual languages can be traced back to innovations in western Proto-Romance. The putative Celtic substrate explanation, which as Posner (1996) explains has at times been suggested as an explanation for the initiation of the lenitions in Romance can be rejected because the phenomena in the two groups of languages, while similar in some respects, have some very different properties, including a different set of inputs and outputs and relevant phonological environments. I turn to Celtic developments in section 3.2.1.8. The similarities, I suggest later, are simply due to the commonness of lenition.

the processes in the present-day synchronic state can contribute to our understanding of lenition.

As is widely described in the literature (eg, Harris 1969, Macpherson 1975, Harris-Northall 1990, Penny 1991, Wireback 1993), a number of processes have been innovated in the history of Spanish which have been described as lenitions - spirantisation, voicing and degemination. - and one of these, as we will see, is still synchronically active in the phonology of the language. It has been suggested (eg, by Martinet 1952, 1955) that these processes need to be considered together in that they are all aspects of a unifying chain shift. This type of position was largely rejected in section 3.1.6, where it was argued that chain analyses do not uncover a phonological, psychological, articulatory or acoustic truth. It seems entirely reasonable to investigate each of these processes individually and we may note, specifically to the Spanish case, that this position fits well with the diachronic facts. After a detailed investigation of possible inter-relationships between the various lenition processes for which there is evidence in the history of Spanish, Wireback (1993) shows that the processes involved in each of these are entirely phonologically different from each other. He concludes: "the various elements of the lenition process - degemination, voicing, spirantization and deletion - are independent processes without a causal relationship" (1993, 340).

In this section I chiefly discuss only one of the types of processes, spirantisation, although a few brief comments on the 'voicing' are appended. As for example Harris-Northall (1990) and Penny (1991) show, a spirantisation process has been innovated more than once in Spanish. Using terminology similar to that encountered here in chapter 2, Harris-Northall (1990) describes a 'first consonant shift' and a 'second consonant shift' in Spanish historical phonology. I focus here on one of the processes in the 'second shift'.⁴⁷ This is a spirantisation process, which, as was mentioned above, is still active in the phonology. It is typically described as a process which had 'voiced stops' as its input and 'voiced fricatives' as its output. Given the recognition of laryngeal realism in section 3.1.3.1, it is important that we do not simply take such statements at face value. Where a language has two series of underlying stops (as in Spanish) and where it is possible, we need to consider whether the language is a

⁴⁷ The details of the interaction between the two shifts are quite complex, but they do not alter anything discussed here. Briefly, the input to the second shift was the output from the first shift (in a similar way to that which was described for the GCS and the HGCS in chapter 2). One of the processes in the second shift deleted /ð, ɣ/ which had been produced by one of the processes of the first shift, but /β/ remained and merged with the output of the second shift which is to be discussed above; see Harris-Northall (1990) for details. I ignore the deletion process here and focus on the other process which formed the part of the 'Second Spanish Consonants Shift'.

|spread| language or a |voice| language.⁴⁸ For Spanish, this is quite easy to ascertain: for example, Macpherson (1975) reports that there is no aspiration in the *Ts* of Spanish and the *Ms* are fully voiced. This clearly qualifies Spanish as a |voice| language with /p⁰, t⁰, k⁰ / : /b, d, g/.

With this in mind, we can give a standard description of the lenitions as in (3.34) which includes both the individual processes (because the lenition is clearly still synchronic, the symbol '→' is used) and some data for environment C (which is taken from Penny 1991); in this data the Spanish orthographic form in fact transcribes the output of the first shift:

(3.34)

b → β	<i>cuba</i>	[kuβa]	'cask'
d → ð	<i>cadena</i>	[kaðena]	'chain'
g → γ	<i>seguru</i>	[seyuru]	'safe'

All of the processes here fit well with the type of lenition trajectory given in section 1.4, although, as we will see shortly, they may not quite be correctly formulated in this traditional version. In any case, the processes described above are affected by the phonological environment of the segments involved. Both prosodic and melodic factors can be seen to be at work. The environments from (3.33) are examined one by one in (3.35):

(3.35)

- A [__#] lenition occurs (although most final consonants had been lost in Spanish before the introduction of this process)
- B [__c] lenition occurs
- C [v__v] lenition occurs (the placement of stress is not relevant: there is lenition in both C1 and C2)
- D [c__] lenition occurs, except not after a nasal, and /d/ does not lenite after /l/
- E [#__] lenition can occur in word initial position: in utterance-initial position, no lenition occurs, but utterance medially, the same generalisations hold concerning preceding melody as in D

Evidence for these generalisations is given in (3.36), where the data has been taken from Martinet (1952), Harris (1969), Macpherson (1975), Harris-Northall (1990), Harris (1997), Butterfield et al. (1997). The data is fully representative. Environment D is exemplified in some detail and the same generalisations hold for environment E which is only exemplified minimally:

⁴⁸ Naturally, we should also consider the possibility that it might be a |constricted| language, but that is normally clear because |constricted| has an easily recognised acoustic effect. Spanish does not feature ejectives (or implosives).

(3.36)

- A *club* [kluβ]⁴⁹ 'club'
Madrid [maðrið] 'Madrid'
- B *obtener* [oβtener] 'obtain'
advertir [aðβertir] 'warn'
signo [siyno] 'sign'
- C1 *sabe* [sáβe] 'knows'
 C2 *saber* [saβér] 'to know'
- D *alba* [alβa] 'dawn' *arbol* [arβol] 'tree'
ambos [ambos] 'both'
caldo [kaldo] 'stock' *abdomen* [aβðomen] 'abdomen'
andar [andar] 'go'
algo [alyo] 'something' *subglottal* [suβɣlotal] 'subglottal'
ganga [ganga] 'bargain'
- E *en Barcelona* [embarθelona] 'in Barcelona'
a Barcelona [aβarθelona] 'to Barcelona'

It is clear from (3.35) and (3.36) that the lenition occurs in most environments, but in certain cases, it does not occur in D and E. The data also shows that there is a potential for a distinction between different 'E' environments: as well as the isolated, utterance initial environment where lenition does not occur, there can be a difference between the effects of the two slightly different environments [c #__] and [v #__]. Lenition occurs even when a consonant is word-initial if it is preceded in a sentence by a word-final vowel; it does not occur if preceded by a word-final consonant of the types illustrated for D.

The above is a quite standard presentation of this relatively well-know Spanish data. However, some important recent work on this phenomenon has cast doubt on the segmental formulae that have been used, following tradition, in this section, for example in (3.34). Lavoie (2000, 2001) has investigated the process described here in contemporary Mexican Spanish in substantial detail, using both instrumental acoustic analysis and electropalatography. An important result is that she actually found no evidence for realisations of /b, d, g/ as fricatives at all. Rather, all her tokens of /b, d, g/ in the relevant environments do not have the narrow

⁴⁹ Other pronunciations are possible, illustrating other processes, but as Harris-Northall (1990) shows, [b] and [d] do not occur.

constriction and acoustic turbulence which are the typical descriptors for fricatives. All the segments which she investigated had much less stricture than would be expected for fricatives and were in fact approximants or glides. The implications of this are that previous descriptions of the phenomenon, including the treatment above, have used symbols such as [β, ð, ɣ] to transcribe approximants. It is better to use unambiguous symbols to indicate the fact that the segments are actually approximants, and from now on I use IPA [β, ɸ, ɥ]. It is not absurd to use the fricative symbols to represent the approximants involved because it avoids the use of a diacritic and it is certainly not the case that the approximants contrast with voiced fricatives, but this may have an impact on our understanding of possible lenition trajectories. We return to this point below, where I use the approximant symbols for clarity. For the moment we can note that these processes do involve ‘opening’ and hence can be seen to fit on the ‘opening trajectory’ of lenition.

This set of data from Spanish spirantisation thus provides solid evidence for potential patterning in lenition and we return to draw out its implications below. Briefly, however, before we leave the Spanish situation, we might consider the diachronic origin of the synchronically leniting stops /b, d, g/. If we compare the original three words used in (3.34) with their Latin etymons, we can observe the situation in (3.37):

(3.37)

Latin	Spanish	
<i>cupa</i>	<i>cuba</i>	‘cask’
<i>catena</i>	<i>cadena</i>	‘chain’
<i>secura</i>	<i>seguro</i>	‘safe’

As Harris-Northall (1990) explains, this illustrates part of the ‘first Spanish consonant shift’. This is a case of ‘sonorising’ lenition in the terminology of section 1.4 and it featured a much less intricate interaction with the phonological environment than does the spirantisation in (3.34). Harris-Northall (1990) claims that it occurred only in environment C [v__v]. If we consider Latin in terms of laryngeal realism, it seems fair to conclude that it was a |voice| language on comparative grounds from the fact that all its daughter languages are |voice| languages. This means that the process involved was as in (3.38):

(3.38)

$p^0 > b$

$t^0 > d$

$k^0 > g$

It is clear from the synchronic analysis of Macpherson (1975), which was mentioned above, that the relevant stops in Present-Day Spanish do indeed include the element [voice]. We now turn to further data from other languages where the treatment will not always be as detailed, often because the processes are less extensive.

3.2.1.3 Proto-Iranian

In quite a different place and different time, lenitions are recorded as diachronic processes in the history of Proto-Iranian, a member of the Indo-Iranian branch of Indo-European. We have already discussed aspects of the phonology of this family of languages in sections 2.3.2 and 3.1.3.1.8. Here we will see that some classic lenition processes have occurred in their history.

The classical description of the historical phonology of Proto-Iranian is Bartholomae (1894).⁵⁰ Together with Old Indic, the language makes up Indo-Iranian, and it is reconstructed from the attested Old Iranian languages such as Avestan (one of the oldest attested Indo-Iranian languages which is the language of the Zoroastrian scriptures, see Bartholomae 1895). In section 3.1.3.1.8, we saw that a series of underlyingly aspirated voiceless stops were formed in Indo-Iranian through fusion with an adjacent [spread] element; this means that the series of segments were /p^h, t^h, k^h/.⁵¹ As Bartholomae (1894) explains, these segments were subject to spirantisation in Proto-Iranian along the lines shown in (3.39):

(3.39)⁵²

$p^h > f^h$

$t^h > \theta^h$

$k^h > x^h$

⁵⁰ See also Beckes (1997) and Sims-Williams (1998a) for sometimes conflicting recent treatments.

⁵¹ As in sections 2.3.2 and 3.1.3.1.8, I ignore certain dorsal complications here and focus on the classic three places of articulation, as specified by [labiality], [coronality] and [dorsality].

⁵² The precise nature of the laryngeal specification of the output is not clear but I assume, following Vaux (1998) that 'voiceless fricatives' are typically specified for [spread]. In any case, the null assumption is that only the manner of the segment has changed; unless there is evidence for a change in the laryngeal specification I assume it remains unchanged.

In terms of the environments in which this process occurred, Bartholomae (1894) reports that it was innovated everywhere except after sibilants and nasals, so it seems that the lenition occurred to a segment in environments A, B, C and E (unless the phonotactics precluded its presence there), but not in specific cases of environment D [c__], which can be represented symbolically as [s__] and [N__]. Comparative evidence (from Bartholomae 1894)⁵³ for the lenition and the environments in which there were exceptions is given in (3.40). This compares the segments in a word from an attested Iranian language (in the third column) with the cognate from Old Indic (in the second column). The input segment is given in the first column, as are the ‘exceptional’ environments:

(3.40)

p ^h	Old Indic	<i>saphas</i>	Avestan	<i>safanho</i>	‘hooves’
[s__]	Old Indic	<i>sphurati</i>	Avestan	<i>frasparat</i>	‘he hurried’
t ^h	Old Indic	<i>yatha</i>	Old Persian	<i>yaθa</i>	‘how’
[s__]	Old Indic	<i>sthanam</i>	Old Persian	<i>stanam</i>	‘place’
[N__]	Old Indic	<i>panthas</i>	Avestan	<i>panta</i>	‘path’
k ^h	Old Indic	<i>kham</i>	Avestan	<i>xa</i>	‘source’
[s__]	Old Indic	<i>skhalati</i>	Avestan	<i>skarayat</i>	‘to make swing’
[N__]	Old Indic	<i>khanati</i>	Avestan	<i>hankanayən</i>	‘they should dig in’

The first line for each of the three segments is representative of all environments other than the exceptional ones described above and in (3.40). Some of the evidence here also shows the effects of certain later changes (eg, deaspiration). I return to the implications of this below.

3.2.1.4 Kannada

The next process is better known in the lenition literature than the Proto-Iranian data just discussed, as it is mentioned in Lass & Anderson (1975) and Lass (1984, 1997) and also in Foulkes (1993, 1997); detailed discussion is found in Tuttle (1930) and Zvelebil (1970). There have been wide waves of lenition in Dravidian languages, including both sonorising and

⁵³ The symbols that Barthomolae (1894) uses have been adapted to the IPA, in line with the discussion in section 2.3.

opening lenitions. I focus here on one very specific aspect of one specific lenition from the history of one of the South Dravidian languages, Kannada, a language with a long-cultivated literary heritage (see Zvelebil 1970). Most of the lenitions that we have seen in this thesis have involved whole classes of segments and it is indeed most frequently the case that a lenition process affects several segments in the same way. This is not always the case, however, and the example here from Kannada is intended to illustrate the case of a lenition process which only affects one segment.

Along with the general lenition processes which have affected Dravidian languages, the Proto-Dravidian segment /p/ underwent a special fate in Kannada. The change that I focus on here is a debuccalisation of the ‘voiceless’ labial plosive to /h/. There seems to be little clear evidence as to the laryngeal state of the original plosive but, in discussing precisely this process, Zvelebil (1970) writes that certain Dravidian scholars assume that that the pre-change plosive was /p^h/; he states “[a] strong aspiration of initial *p*-, common to all Tamil dialects spoken in Ceylon would point in this direction” Zvelebil (1970, 86).⁵⁴ In addition, Lass (1984, 1997) and Foulkes (1993, 1997) argue convincingly that an intermediate fricative stage was involved, although this has since lenited further through debuccalisation. I thus reconstruct the lenition trajectory in question here as in (3.41), which also shows a further possible lenition to zero:

$$(3.41) \quad p^h > f^h > h > \emptyset$$

There is much less to note here in terms of phonological environment, largely because of the general phonotactics of Dravidian. In environments A [__#] and B [__c] there was no lenition, but this is only because Kannada had no consonants in these positions (see Zvelebil 1970). To further complicate matters, we cannot consider environment C [v__v] either, because no labial stop occurred in that environment, in common with all Dravidian languages, because it had been lost in the Proto-language (see again Zvelebil 1970).

The two remaining environments are relevant, because the input did occur there in Proto-Dravidian. In environment E [#__], lenition occurred and the debuccalisation stage is clearly attested in writing; in non-Brahmin dialects, the lenition is ‘complete’ and /h/ has deleted. This

⁵⁴ He comments further that Kannada had borrowed voiceless aspirates from Indo-Aryan languages of the Indo-Iranian family, which would seem to back up the idea that such stops were aspirated in the language.

is illustrated by the data in (3.42)⁵⁵ which compares Kannada with Tamil, a non-debuccalising Dravidian dialect:

(3.42)

Tamil <i>pol</i>	Kannada <i>hol, ol</i>	'resemble'
Tamil <i>puku</i>	Kannada <i>hugu, ogu</i>	'reach, enter'

In environment D [c__], debuccalisation does not occur. In gemination, often no lenition at all occurs and in the environment [N__], which is the only other phonotactic possibility, lenition to /h/ also does not occur; it could be argued that a 'lesser' degree of lenition occurs in that the segment is typically written as an *M*, but it is clear that neither debuccalisation nor spirantisation has occurred here, as shown in the data in (3.43):

(3.43)

Tamil <i>appan, appu</i>	Kannada <i>appa, apa</i>	'father'
Tamil <i>kuppam</i>	Kannada <i>kuppe</i>	'multitude, heap, pile'
Tamil <i>ampu</i>	Kannada <i>ambu</i>	'arrow'
Tamil <i>kompū</i>	Kannada <i>kombu</i>	'branch, horn'

Although only one segment is involved, the pattern of the process fits with others that have been discussed in this thesis and elsewhere as cases of lenition. Also, as will become clear, the interaction of the process with the phonological environment is typical of that found in lenition processes.

3.2.1.5 Celtic

In section 1.4 it was shown that certain processes in Celtic have a special place in the lenition literature, because Thurneysen originally invented the very term 'lenition' itself to describe them. As was also mentioned in that section, one of the ways in which the term 'lenition' is currently used is to describe the morphophonological remnants of past phonological lenition processes which play an important role in the grammar of contemporary Celtic languages. Given all this, it might seem odd not to mention the Celtic situation at all. In this section, therefore, I examine certain aspects of lenition in Celtic. It would go far beyond the bounds of a small section like this to deal with all the details, but certain non-trivial aspects can be extracted from the whole. Substantial discussion of the processes and of issues that are related to them can be found in, for

⁵⁵ The data is from Zvelebil (1970), Tuttle (1930) and Lass (1997).

example, Thurneysen (1909), Pedersen (1909), Morris-Jones (1913), Martinet (1952, 1955), Ball & Müller (1992) and Russell (1995). In the discussion here I only deal with points which are relevant to obstruents, in keeping with the theme of this thesis.⁵⁶

The whole story involves a set of somewhat diverse processes, all of which would reasonably fit on lenition trajectories of the type given in section 1.4. These processes were innovated in insular Celtic languages,⁵⁷ doubtless partly as shared innovations, and have left a clear mark on the languages in the form of the morphophonological alternations just mentioned. These can involve most consonantal segments when they occur in word-initial position. The phonological processes could be diachronically captured by the grammar because the lenitions occurred, rather like in the Spanish data from section 3.2.1.2, in environment E [#__], to the extent that this was [v #__], within certain types of syntactic constituents. The phonological lenition processes also occurred in other environments (eg, C [v__v]), but the segments that were produced here were later lexicalised as underlying segments and the only alternations that survived were those which alternated in initial positions, depending on the syntactic and lexical context in which they were placed. The conditioning vowels which were originally word-final have since often been lost, but the morphophonological alternations remain, now entirely divorced from phonological conditioning.

These alternations are clearly no longer motivated by a phonological environment and now either serve to indicate grammatical function (eg case, gender, direct-objecthood) or are an automatic reflex of the conjunction of particular words. The modern morphophonological processes are often referred to as (initial) mutations⁵⁸ and the precise inventory of them varies quite substantially among the Celtic languages (and within them dialectally). The precise nature of the synchronic processes are quite different in members of the two surviving branches of Celtic (Brythonic and Goidelic, for example Welsh and Irish, respectively); this is a good indication that the original phonological processes which led to them were also different.

⁵⁶ Sonorants are involved in some of the synchronic morphophonological processes; for example, in Welsh, the alternations are: p ~ b ~ f ~ m; t ~ d ~ θ ~ n; k ~ g ~ x ~ h; b ~ v ~ m; d ~ ð ~ n; g ~ Ø ~ ŋ; m ~ v; t ~ l; r ~ r.

⁵⁷ It is not entirely clear whether continental Celtic languages such as Gaulish shared any aspects of lenition with the Insular languages given the paucity of evidence that we have for them (see, for example, Russell 1995, Sims-Williams 1998b).

⁵⁸ These synchronic morphophonological processes are also at times referred to as 'lenition', especially in discussions of the grammar of Irish, but it should be noted that this usage gives the terms a very different meaning to that which it has in this thesis; the two meanings are obviously related in that phonological lenition (my interpretation) gave rise to morphophonological lenition, but their ontological status is different.

The interest here is naturally on the synchronic phonological processes which once were innovated in these languages. As for example Thurneysen (1909) and Russell (1995) explain, these processes included ‘degemination’, ‘voicing’ and ‘spirantisation’ and some have been tempted to formulate chain analyses which link them (eg, Martinet 1952, 1955). As above however, it is not clear that this can relate to the causation of the processes and, as Russell (1995) writes, recent work such as Thomas (1990) and Sims-Williams (1990), shows the processes to be separable.

I turn here to the most detailed description of some of these processes: Thurneysen’s (1909) account of the history of Irish, where he focuses on spirantisation of the underlying stops. We may well wonder whether Pre-Irish was a |spread| language or a |voice| language. It is clear that the Present-Day languages are |spread| languages (see Ball & Müller’s 1993 report on instrumental measurements for Welsh, for example).⁵⁹ Koch (1990) and Martinet (1952, 1955) both project this back for various reasons to pre-lenition Celtic, and I follow them in assuming that the language was a |spread| language. Thurneysen (1909) explains that the spirantisation affected both series of stops *except* in the sets of environments given in (3.44), translated directly from Thurneysen (1909, 70-71), retaining his symbols for simplicity:

(3.44)

- a) all double (lengthened) consonants, even when they lost their length when adjacent to other consonants...
- b) the old clusters *ng* (ie, *ŋg*), *nd*, *mb*, *sc*, *st*, (*sp* in loanwords)...
- c) the stops after *r* and *l*; *t* after *ch*; *b g* after *ð* which developed from *z*⁶⁰

The cluster *sp* (ie, /sp^h/) only occurred in loan words because Irish has completely lost etymological /p^h/, in a similar way to that reported for Kannada in section 3.2.1.5. While I ignore the exceptions for non-obstruents, we can see that the environment D [c__] has an important effect, including geminates. There seems to be an interesting range of melodic effects

⁵⁹ They write: “Traditionally, the plosives of Welsh have been described as consisting of two sets: /p, t, k/, termed voiceless (or fortis), and /b, d, g/, termed voiced (or lenis). However, the spectrograms produced showed that very little voicing (if any) occurred during the closure stage for /b, d, g/. An alternative description might be (as in English) to utilize the terms ‘aspirated’ and ‘unaspirated’ instead of voiceless and voiced respectively.” (1992, 84).

⁶⁰ The original reads: “a) alle doppelten (gedehnten) Konsonanten, auch dann, wenn sie unmittelbar neben andern Konsonanten ihre Dehnung einbüßten [...] b) die alten Gruppen *ng* (d.i. *ŋg*), *nd*, *mb*, *sc*, *st* (*sp* in Lehnwörtern) [...] c) die Verschlusslaute nach *r* und *l*; *t* nach *ch*; *b g* nach dem aus *z* entstanden *ð*”

here, where the precise nature of the ‘c’ is important. We return to these points in 3.2.3 and turn now to some further lenitions.

3.2.1.6 Greek

All of the ‘opening’ lenitions that we have investigated thus far have shown some environmental effects. It seems, however, that such lenitions can occur with no reference to environment at all. The case of Greek *TAs* is a quite well-known case where all occurrences of the input segments are affected by a process. Once the process became non-variable, this naturally led to the complete loss of the type of segment concerned from the system. As Sturtevant (1940) and Palmer (1980) show (and see also, for example, Buck 1933 and Horrocks 1997), we can be quite sure on philological evidence that Ancient Greek had a stop series comprising /p^h, t^h, k^h/; these segments derive from IE *MA*s and were written ⟨φ, θ, χ⟩. From around the 4th century BCE, up to certain cases in the first century CE, spellings appear in manuscripts to show the innovation of a process which has left these segments as fricatives in Present-Day Greek (see Sturtevant 1940; the earlier date refers to the innovation in only certain dialects). There is, however, philological evidence for an intermediate affricate stage; for example there are spellings such as ⟨κχ⟩ for etymological /k^h/ at early stages of the change so it seems likely that this process involved two quanta (see the discussion in section 12.2.3 - Horrocks 1997 assumes this to have been the case, and see the comments in Sturtevant 1940). On the assumption that there were two quanta and that it did not affect the laryngeal specification in the segments,⁶¹ the process can be represented as in (3.45):

$$(3.45) \quad \begin{array}{l} p^h > pf^h > f^h \\ t^h > t\theta^h > \theta^h \\ k^h > kx^h > x^h \end{array}$$

While all word-final stops had previously been lost (see Palmer 1980), it seems that all other remaining occurrences of /p^h, t^h, k^h/ in Greek underwent the processes and there was neither a prosodic nor melodic interaction with the environment.

⁶¹ Nothing substantial rests on this decision, but see the comments on this issue in 3.2.1.3.

3.2.1.7 Bantu

Perhaps even more than in Dravidian (see section 3.2.1.4) and certain other language families that we have discussed, there have been substantial lenitions in Bantu languages. The treatment here will be far too brief to detail the precise nature of the various processes in each language, and the section is partly included to illustrate the fact that analogous processes to those which we have seen in Indo-European and Dravidian languages have also occurred in a language family which is entirely geographically and genetically unconnected.⁶²

In a discussion on the classification of Bantu languages, Nurse (1999) describes 'Bantu Spirantisation' and writes that "...this covers a huge area and has affected many Bantu languages [...] Most of the languages not affected are contiguous and spoken in the rain forest ... With very few exceptions, all Bantu languages south and east of the rain forest have been affected ..." (1999, 21). Classical discussion of Bantu spirantisation is found in Meinhof (1899) and more recent work which touches on the topic includes Meeussen (1967) and Guthrie (1967-1971).

Meinhof (1899) assumes that Proto-Bantu had only one series of stops: /p, t, k/⁶³ and goes on to illustrate a range of reflexes in different languages, including /ɸ, s, x/ in Sotho⁶⁴ and /β, h, k'/ in Digo and /p, t^h, x/ in Sango, all of which illustrate at least some lenition. Subsequent work has shown (see Meeussen 1967 and Hyman 1999) that what Meinhof assumed was a series of fricatives in proto-Bantu was also originally a stop series /b, d, g/. Doke (1954) and Tucker & Bryan (1957) supply correspondences for these segments from a large number of Bantu languages, including /β, l, j/ for Northern Sotho, /v, l-r, ɣ/ for Taita and /Ø, r, ɣ/ in Kikuyu. These processes illustrate a range of lenitions (including processes with non-obstruent outputs which we have focused on here).

There was little phonotactic freedom in Proto-Bantu, as Meinhof (1899) and others have shown. This means that few of the environments given in (3.30) can be discussed, but there were

⁶² The possibility exists, of course, that Bantu, Indo-European and Dravidian, along with all other language families are ultimately related, and proponents of 'long-range' reconstruction seek to reconstruct both proto-languages and the relationships that exist between universally recognised language families (see, for example, Bomhard 1986 and Ruhlen 1987). It is not clear that they can be successful, however (see Salmons 1992, Fox 1995, Lass 1997), because much of the heart of the comparative method is rejected, as was noted in chapter 2. The point about the difference that exists between language families holds, in any case.

⁶³ A palatal stop is also reconstructed, but as palatals often undergo assibilation, I leave them out from consideration here; this point applies throughout the section.

⁶⁴ Meinhof's (1899) transcription is different, but he provides a clear key which allows translation to the IPA; for example, the segment that I transcribe here as /s/ is given as /r/, but this is glossed as 'fortis voiceless alveolar fricative'.

clusters of nasals and stops, in common with many languages which otherwise do not allow clusters (these are often referred to as ‘Prince’ languages, see Prince 1984, Harris 1997). While stops elsewhere have often lenited in Bantu languages, the stops in this [N__] environment are often not affected. I give one example of several, from Tharaka, taken from Tucker & Bryan (1957) with Proto-Bantu consonantism from Meeussen (1967). The full correspondences (excluding the palatals) for this language, including nasal + stop clusters, are shown in (3.46):

(3.46)

Proto-Bantu		Tharaka
p, t, k	>	h, t, k
b, d, g	>	Ø, r, ɣ-g
mp, nt, ŋk	>	mp, nt, ŋk
mb, nd, ŋg	>	mb, nd, ŋg

While the discussion in this section has been somewhat oversimplificatory, what is clear is that many Bantu languages have initiated lenition processes, just like many languages from other genetic groups and geographic areas; also while prosody does not seem to have any effect, the lenitions often can be seen to interact with their melodic environment.

3.2.1.8 Liverpool English

The last set of lenition data that I discuss here shares certain substantial similarities with many of those already mentioned. It is a synchronic process like the Spanish data in section 3.2.1.8, it shows certain clear environmental effects like many of those discussed here and it features overt affrication and substantial spirantisation, like the HGCS₁. Unlike other data discussed here, however, the process is subject to sociolinguistic variation. It is one of the classic defining features of the accent of English which is centred around the city of Liverpool, England. This variety is also spoken in the surrounding conurbation of Merseyside, but it is quite tightly localised for various geographic and historical reasons. The variety is often seen as stigmatised in Britain (although this prejudice does not seem to extend far outside the British Isles) but it has substantial covert prestige and seems to be spreading.

The serious study of Liverpool English can be traced back to Knowles (1974), which provides a general overview of the variety and substantial detail on several phonological points, especially on the suprasegmental level, as well as some comments on the history of the variety. Knowles

discusses segmental phonetics in some detail, but the phonological patterning of features is not treated in depth. Subsequent work on the variety consists principally of Knowles (1978), de Lyon (1981), Newbrook (1986 and 1999), Sangster (1999, to appear) and Honeybone (2000d, 2001). Discussion is also found in Harris (1990 and 1994) and general volumes on English accents such as Wells (1982) and Hughes & Trudgill (1996). While Knowles (1974) and the other studies listed above describe many of the basic features of the variety, the detail of many of these features is still not fully known. In this section I report the results of some preliminary and ongoing investigation of the patterns that can be seen in the lenition which characterises the area (see Honeybone 2000, 2001 for further details and methodology).

Liverpool English is clearly a *|spread|* language. When the *T*s are not affricated or spirantised, they feature definite aspiration and the *M*s are also the same as in reference accents (such as RP and General American) in this regard. The lenition processes can affect all stops, but are most frequent in */t^h, k^h/* among the *T*s. I restrict my comments here to these two segments, as these are the best researched.⁶⁵ The lenitions are not well known outside of the Accent Studies literature (for the notion ‘Accent Studies’ see Foulkes & Docherty 1999), although certain aspects of them are briefly mentioned in Labov (1994, 2001), Trask (1996), Hock & Joseph (1996).

The typical lenition processes are summarised in (3.47). The information given there illustrates the fact that two stages of lenition can be observed (variably) for both segments in some environments; they can be thought of as synchronic lenition trajectories for which all stages can be observed in the present-day linguistic system:

(3.47)

$$\begin{array}{l} t^h \rightarrow t\underline{\theta}^h \rightarrow \underline{\theta}^h \\ k^h \rightarrow k\underline{x}^h \rightarrow \underline{x}^h \end{array}$$

As we will see below, there is clear environmental interaction in these lenitions and they are most easily conceived of as stages of lenition, such that in certain environments, no lenition is possible, in others, lenition to an affricate is possible, and in still others, lenition to an affricate or fricative is possible. A few notes are in order in connection with (3.47):

⁶⁵ It seems clear that the basic generalisations here also extend to */p^h/*, although likely with less frequency; the processes are variable and it is quite possible that speakers vary as to the frequency with which they lenite the different stops. As Sangster (1999, to appear) reports, similar effects can be observed for */d/*.

(i) just as in section 2.3.1, I ignore the precise place of the |dorsal| fricative and retreat to a standard transcription [x]. The reasons behind this decision are partly the same as in section 2.1.3 and partly slightly different. As with the ‘Germanic’ dorsal fricative, the ‘default’ place of articulation seems to be uvular (at least pre-uvular), but the tradition is to transcribe it as [x]. The decision is therefore partly a matter of tradition. Also however, the decision is a conscious simplification to avoid the need to comment on a process of place assimilation. While the default place of the fricative may be uvular, it can undergo assimilation to [ç] in a front high vocalic environment. The details of this are complex however and would take us too far afield from the topic of lenition (they are similar, but not quite the same as those which underlie the alternation between [x] and [ç] in German). The place of this lenition-produced fricative is consistently dorsal and I adopt the symbol [x] as a cover-symbol for ‘dorsal |spread| fricative’.

(ii) for some speakers the |coronal| fricative and affricate have a slightly sibilant quality which makes them closer to canonical grooved [s], but there is quite consistently a distinction between etymological /s^h/ and the fricative lenition output of etymological /t^h/, along the lines explained in section 2.3.1; the slit fricative [θ] is the typical output (see Honeybone 2001 for some further discussion). The maintenance of contrast between these two segments along with the distinctive nature of the characteristic lenition output of /t^h/ are the basis for the proposed analoguehood of this situation with the lenition of the |coronal| stop in the HGCS₁. The fact that the lenition processes and environments are also practically identical, as will be seen shortly, are also encouraging in this regard.

(iii) there is also a further stage of lenition for /t^h/ in final position; the segment debuccalises to [h] in a very small set of environments which are at least partially characterisable lexically. It is likely that this is a lexically diffusing process which may be spreading to take in other environments. I leave the investigation of this to future research.

To illustrate the environmental interaction, I return in (3.48) to the method of presentation used for the Spanish data. This is particularly applicable to a linguistic system such as Liverpool English because it has quite a wide phonotactic freedom. This assumes the way of understanding the processes which was outlined above and acknowledges that the processes are variable: (i) where I claim that ‘lenition to a fricative is possible’, lenition is also possible to an affricate, as is no lenition at all; (ii) where I claim that ‘lenition to an affricate is possible’, this precludes lenition to a fricative, although stop realisations also occur; (iii) where I claim

that ‘no lenition is possible’ the typical surface form of the underlying stops is a surface stop.⁶⁶ In general, underlying /t^h/ can be realised as an affricate in most environments and as a coronal fricative in a subset of these environments. Underlying /k^h/ can be realised in most environments as a dorsal affricate and in a subset of these it can be further lenited to be realised as a dorsal fricative:

(3.48)

- | | |
|-------------------|--|
| A [__#] | lenition is possible to a fricative for both /t ^h / and /k ^h / |
| B [__c] | lenition is possible to a fricative for /k ^h /; the segment /t ^h / hardly occurs in this environment due to phonotactic constraints |
| C1 [ˇ__(v)] | lenition is possible to a fricative for both /t ^h / and /k ^h / |
| C2 [(v)__(ˇ)] | lenition is possible to an affricate for both /t ^h / and /k ^h /; it is possible that lenition to a fricative is also possible here for /k ^h / |
| D [c__] | the degree of lenition varies according to the melodic content of ‘c’; see below in (3.46) |
| E [#__] | lenition is possible to an affricate for both /t ^h / and /k ^h / |

The environment D [c__] requires special comment, as it so frequently has in the discussion in this section. The generalisations are slightly different for the two underlying segments. They are summarised in bullet-pointed form in (3.49):

(3.49)

- in [s__], there seems generally to be no lenition at all for either segment
- in [N__], only lenition to affricates occurs (to the exclusion of fricatives) for both /t^h/ and for /k^h/
- for /t^h/, the environment [l__] only allows lenition to affricates, whereas [1__] allows lenition to fricatives for /k^h/
- it seems further that [t^h__] and [k^h__] inhibit lenition for /t^h/

Evidence for the processes is given in (3.50). This consists of a list of words from Liverpool English which illustrate the most lenition possible for the environments in (3.48) and (3.49). The data is largely taken from Honeybone (2000, 2001). As the symbols are ‘surface’ phonetic symbols, traditional transcriptions are used, but they are to be thought of as |spread| segments.

⁶⁶ As Docherty, Foulkes, Milroy, Milroy, & Walshaw (1997) show, more is ‘possible’ in phonetics than phonology normally predicts, so it may be that even where I claim that ‘no lenition is possible’, non-stop surface forms occur, but these are the same type of non-stop surface form that occur in all linguistic systems due to variance in articulation and are not the characteristic ‘Liverpool lenited’ forms (see Honeybone (2001) for further discussion).

(3.50)

A [__#]	<i>alright</i> [θ]	<i>book</i> [x]	
B [__c]	---	<i>respect</i> [x]	
		<i>basicly</i> [x]	'basically'
C1 [\acute{v} __ (v)]	<i>city</i> [θ]	<i>crack<u>ers</u></i> [x]	'mad'
C2 [(v) __ \acute{v}]	<i>attack</i> [tθ]	<i>okay</i> [kx]	
D [c__]	<i>fifteen</i> [t]		
	<i>respect</i> [t]		
[s__]	<i>station</i> [t]	<i>scally</i> [k]	a derogatory term
[l__]	<i>adult</i> [tθ]	<i>wel<u>come</u></i> [x]	
[N__]	<i>moment</i> [tθ]	<i>in<u>con</u>venience</i> [kx]	
E [#__]	<i>taken</i> [tθ]	<i>come</i> [kx]	

While it is an abstraction from sociolinguistic variability, the data here shows the effect that prosodic and melodic factors have on the kind of opening lenition that is found in Liverpool English. Further work on the phenomena is pressingly needed, but I believe that generalisations in this section are robust.

3.2.1.9 Summary and prospect

The lenitions covered in this chapter have been quite diverse. While they have included some 'voicing', most involved 'opening' of one sort or another. Many, but not all of the lenition processes have been the subject of some kind of environmental effect, but we have seen that this can be quite diverse. The heuristic environments formulated in (3.30) have proved to be useful for the description of some lenitions, but, if it was hoped that the interaction between process and environment would turn out to be straightforward, that hope has been dashed.

In the next section, I scrutinise the processes which have been identified as 'lenition processes' and consider how each process can be formalised and to what extent all obstruent lenition processes can be unified, using the data presented here and in chapter 2 as a yardstick. After that, in section 3.2.3, I return to the connection between lenition processes and phonological environments and endeavour to make sense of the effects that were shown to hold between the two in this section and in chapter 2.

3.2.2 What is lenition?

We have now considered quite a range of lenition processes, some in fine detail in section 2.1 and others, in section 3.2.1, with at least enough information to understand their general patterning. As is clear from the structure and avowed aims of this thesis, this imbalance in presentation is intentional. The philological and, latterly, linguistic detail that was expended on the presentation of the processes, or sets of processes, in chapter 2 was at least in part intended as an investigation into what precisely we (think we) know about these processes in general and about the situation which surrounded their innovation in the particular languages involved. I argue later that, despite their sometimes iconic nature in historical phonology, we have not quite yet understood them as individual events, largely because they have not been considered in a cross-linguistic lenition and laryngeal realist perspective. In chapter 4, I reinterpret these four phenomena in terms of the model of phonology and diachronic innovation defended here. All discussion of the details associated with those cases will thus be encountered in a later chapter when the philological concerns of what went on precisely in specific languages re-emerge into the discussion. Many of the types of processes that were involved in the GCS, HGCS, EIFV and IGCW will feature in the generalised analysis of lenition processes that is to occur in this section, however, so at least part of their analysis will be effected here. Chapter 4 brings the details of the general discussion back down to the specific, and deals with the issues raised by the innovation into actual linguistic systems of the types of processes which are discussed in this section.

Given the data in section 3.2.1 and chapter 2 (along with the laryngeal realist discussion of the laryngeal phonology of modern Germanic languages and of Proto-Germanic in section 3.1.3.1), there are several quite fundamental observations that we can make about attested lenition processes and some key generalisations about the phonological patterning that they show. This section is built around these observations and generalisations. In the first subsection (3.2.2.1), they are set out in a simple form. In sections 3.2.2.2.1 - 3.2.2.2.6, I consider certain aspects of previous phonological treatments of lenition and weigh them up, partly as theoretical wholes, and also against the points recognised in 3.2.2.1, in order to ascertain how insightful such previous approaches are and what contributions they can be thought to make to 'lenition theory'. In section 3.2.2.3, I conclude this section by reconsidering the processes involved and by proposing a novel understanding of lenition, which, I argue, is the only way to coherently consider the concept.

There are also certain clear generalisations which can be made about the interaction between lenition and phonological environment, but we shall investigate those in the next section (3.2.3). We focus here on the segmental processes involved.

3.2.2.1 Lenition observations and generalisations

The observations and generalisations that can be made concerning lenition can perhaps best be seen in connection with the general set of lenition processes which, when put together, make up lenition trajectories. The pre-laryngeal-realism lenition processes recognised in section 1.4 are repeated here as (3.51). I focus on these one by one in this section:

(3.51) Lenition processes

affrication: $t \Rightarrow ts$, $k \Rightarrow kx$

spirantisation: $k \Rightarrow x$, $d \Rightarrow \delta$

voicing: $f \Rightarrow v$, $p \Rightarrow b$

debuccalisation: $s \Rightarrow h$, $x \Rightarrow h$

From the cases of affrication that we have considered in this thesis, one generalisation is clear: there is a correlation between the presence of [spread] in stops and affrication. All the cases of affrication that have been considered here occur in stops which have [spread] as their (only) laryngeal element. The obvious cases of this are that of Liverpool English and the HGCS₁ (given the recognition of the laryngeal phonology of Germanic in section 3.1.3.1). Additionally, as was mentioned in section 3.2.1.6, if we accept the philological evidence that the Greek process involved two quanta, then all three processes fit a general schema of $k^h \Rightarrow kx^h \Rightarrow x^h$ (exemplified for velars).⁶⁷

It is tempting to seek a causal relationship between these two points: a reasonable assumption, given the [spread] input and the affricate output is that [spread] in some way enables a lenition process of this type to occur.⁶⁸ What is needed to cement the causal connection is a phonetic or phonological mechanism through which the process can be innovated into the phonology of a language, and I consider the ability of previous accounts of

⁶⁷ Without evidence to the contrary, we can assume that the affricates retain the [spread] of the input. This keeps down the number of quanta involved in the process and seems clearly right for the affricates of both Present-Day German and Liverpool English.

⁶⁸ Given the discussion of the notions of 'causation' and 'explanation' in section 1.3.1, it is only with reservations that we can write that "[spread] causes affrication", although this would be the normal way of expressing the observation.

lenition to provide this in the coming sections of this chapter. If we take the three suggested cases of affrication, it seems that (i) in the Liverpool case, a synchronic process of affrication has been retained, along with spirantisation in certain environments, that (ii) in High German the affricates have been lexicalised in certain environments and fricatives have been lexicalised in other environments,⁶⁹ and that (iii) in Greek, a further lenition to fricatives has occurred across the board.

Given the cases of spirantisation which we have encountered in this thesis, it seems that there are, in fact, at least two types: (i) cases which involve *|spread|* in the input and behave in a very similar way to the cases of affrication just discussed (especially for the Greek case where affricates have been entirely lost); these include the Proto-Iranian case and, it seems likely, the fate of */p^h, t^h, k^h/* in Irish, and GCS₁ (if we adopt the version in (2.35) which follows the Glottalic Theory) which seem to involve */p^h, t^h, k^h/* \Rightarrow */f^h, θ^h, x^h/*. There is also: (ii) the Spanish type, which involves *|voice|* stops and produces *|voice|* approximants (ie, */b, d, g/* \Rightarrow */β, ð, ɣ/*), if we follow Lavoie (2000, 2001); this may well also include the GCS₂ (we return to this in chapter 4, of course). This may be similar to the case of the *Ms* in Irish and Bantu, although these may represent a theoretically possible third type of spirantisation, which involves neutral stops.

There is now clearly a distinction between cases of ‘voicing’. These include (i) a possible */p^o, t^o, k^o/* \Rightarrow */b, d, g/* or */f^o, θ^o, x^o/* \Rightarrow */v, ð, ɣ/*, which we encountered in the history of Spanish in (3.35), and (ii) */p^h, t^h, k^h/* \Rightarrow */p^o, b^o, k^o/* or */f^h, θ^h, x^h/* \Rightarrow */f^o, θ^o, x^o/*, which we have not unambiguously encountered yet, but will do later (the possibility of such a process is clear and will suffice for the moment).

The cases of debuccalisation that we have encountered allow for certain clear observations: there seems to be a correlation between segments which involve *|spread|* and debuccalisation to *[h]*. This is the case for the ‘GCS debuccalisation’ which we encountered in section 2.1.1.5 and 3.1.3.1.8, and for the example from Kannada in section 3.2.1.4, and also for the debuccalisation found in Liverpool English (mentioned briefly in section 3.2.1.8). It also seems to be the case that the direct historical input was a fricative for all of these cases. We also encountered the debuccalisation of ejectives in section 3.1.3.1.8, which results in *[ʔ]*.

⁶⁹ I address the issue of the geminacy of the fricatives in section 4.1.2; for the moment I ignore it (in a way that will be shown to be principled in 4.1.2).

We might hope that any full account of obstruent lenition (any ‘lenition theory’) would consider the above observations and generalisations in some sense. There are several ways in which this could be done, however, and it is not obviously the case that all of these observations need to be accounted for in the same way. A platonically ‘perfect’ lenition theory would account for everything dealt with here, and give an exact and insightful account of all lenition processes as a unified type of process, and would make predictions about what kind of processes can occur in particular environments. We may well wonder, however, whether we can hope for perfection to exist in the mortal world.

After a preliminary discussion in the preamble to the next section, which returns to the notions of prediction and explanation, I turn to several previous proposals and types of analysis which could be conceived of as ‘lenition theories’ in order to investigate how they measure up against the platonic ideal. After this, finding them all wanting in some regard, I propose what I believe to be a coherent understanding of the concept, which comes as close to the ideal ‘lenition theory’ as possible. It does not fulfil all the criteria of the platonic ideal just mentioned, but, as I explain, this is for very good reasons.

We turn now, in the next sections, to a range of previous accounts for lenition. The key question to consider there in connection with the processes described in this thesis and summarised above is: if these are lenitions, then what is lenition?

3.2.2.2 Previous accounts of lenition

As was noted in section 1.4, there have been many previous accounts of lenition, too many, in fact, to consider them all. In this section, I consider a number of these accounts and assess how they answer the question ‘what is lenition?’ I consider whether they can account for the specific observations and generalisations described above and in general I assess how compelling they are and to what extent they can contribute to our understanding of the phenomena. In section 1.3.1, we considered the question of how explanatory the analyses of historical phonology can be. We saw that absolute predictions as to when and where specific lenitions might occur are certainly not to be expected from such analyses, but a certain degree of prediction is to be hoped for, perhaps at least in terms of predicting what will never occur as a natural endogenous innovation or in terms of which set of conditions are necessary for a particular innovation, else it cannot occur. As we also saw, it is possible that certain analyses predict what is *likely* to occur, although it is still not clear to what extent such claims can be seen to be true predictions or in any way explanatory.

It is quite straightforward to give an account of the lenitions that we have considered in terms of phonological rules of the type that were prevalent in the days of Standard Generative Phonology but which few phonologists use nowadays, but this approach will not fare well in terms of accounting for the observations of the last section. This is not so much a failure of any particular account of lenition, but rather, substantive issues arise concerning the explanatory aims of this way of conceiving of historical phonology, as we saw in section 1.3.1.

For example, Lehmann (1973) formulates the rule in (3.52) to account for the fricative outputs of the HGCS₁ (another rule is provided to account for the affricates):

(3.52)

$$\left[\begin{array}{c} + \text{cns} \\ + \text{tense} \end{array} \right] \rightarrow [+cont] / [+voc] _ \left\{ \begin{array}{c} [+voc] \\ \# \end{array} \right\}$$

Now, there are several problems that are associated with rules such as this from the phonological perspective that has been developed in this thesis. One is simply the choice of features, but that is not a substantive problem and the rule could be reformulated to use any desired feature (or element) set and [+voc] could be replaced by a syllabic understanding of the notion ‘vowel’.⁷⁰ Another problem is that it doesn’t quite work for the data because fricatives were produced in a slightly wider environment in most varieties (for example between /l/ and a vowel, as in *helfen* ‘help’), but this, too, can be rectified in a rule of this sort.

A more clearly successful rule of this sort might be the account of GCS₂ which is on the dust jacket of King (1969), as shown in (3.53):

(3.53)

$$\left[\begin{array}{c} + \text{obs} \\ + \text{tense} \end{array} \right] \rightarrow [+cnt]$$

Because this process is assumed to be context free, the rule seems quite compellingly simple, but the problem is the same. Such rules as these do not engage with the kind of observations and generalisations that were made in section 3.3.2.2. As we saw in section 1.2, such rules are not intended to engage with or reflect the phonetic or phonological factors

⁷⁰ It is interesting and maybe even reassuring to note that Lehmann uses [±tense] to characterise the Germanic stops and not [±voice].

which led to the innovation of the process (its ‘cause’). As such, they do not enter the field as possible explanations of lenition and have no connection with possible predictions as to what can and cannot occur.⁷¹ In fact, they can function as clear and useful descriptions of the processes, but not as explanations, a point which has been made quite frequently (see, for example, McMahon 1994, 2000b). While we can model lenition as the introduction of a phonological rule, this does not tell us much about what lenition is, what factors lead to it occurring and what might make a process recognisable as a case of lenition, rather than metathesis, for example.

Similar points apply to certain models of phonological change in OT. For example, Jacobs (1994) analyses certain lenitions in the History of French (which were rather similar to those examined here for Spanish in section 3.2.1.2) using OT constraints and constraint ranking as in (3.54).

(3.54)

- (i) PARSE (voice) >> PARSE (cont)/[-voice] & PARSE (cont)/[+voice] >> *LE/[+voice, -cont] >> *LE/[-voice]
- (ii) PARSE (voice) >> *LE/[+voice, -cont] >> PARSE (cont)/[-voice] & PARSE (cont)/[+voice] >> *LE/[-voice]
- (iii) PARSE (cont)/[-voice] >> *LE/[-voice] >> PARSE (voice) >> *LE/[+voice, -cont] >> PARSE (cont)/[+voice]

We need not go into the details of the interpretation of the constraints, because the point here is that the innovation of lenition is accounted for as a change in constraint ranking.⁷² Two such changes are shown here: from stage (i) to (ii), which represents a spirantisation of voiced stops to voiced fricatives in a specific ‘lenition environment’ (‘LE’), and from (ii) to (iii), which represents the subsequent voicing of voiceless obstruents. In both these cases, constraints have moved on the hierarchy and this makes overt claims as to what has occurred. It does not seek to predict what can and cannot occur, however, nor to account for the type of generalisations which were recognised in section 3.3.2.1. Both of these models, constraint reranking and rule addition, are perfectly fine descriptions of diachronic occurrences, if the model of the grammar which they are associated with is accepted, but in and of themselves,

⁷¹ As is well known, any imaginable process can be formulated using such rules, including ones which never occur and this was one of the key reasons why many phonologists sought to ‘constrain the model’ in the 60s and 70s so that rules reflected naturalness in some way (see, for example, Anderson 1985, Carr 1993 and the discussion in section 1.2.2.2).

⁷² Some of the conventions involved are: ‘>>’ indicates strict domination of one constraint above another, ‘&’ indicates equal ranking, and ‘LE’ refers to a specific ‘lenition environment’ which Jacobs formulates separately to involve intervocalicity and the like. The PARSE command could be replaced by MAX or IDENT to bring this into line with McCarthy & Prince’s (1985) correspondence theory.

they tell us nothing about lenition. It might thus be too hasty to accept such accounts as being all that we can say about lenition processes.

The accounts which I consider here in the subsections of this section generally try to bring our understanding of lenition to a level further than just describing the effects that it has. Some of them make predictions as to what can be taken to count as a case of lenition and also about the kinds of factors which allow for or govern its introduction into the phonology of a language. This gives these approaches a clear theoretical advantage over others because they can be proved wrong, and this makes them interesting and important proposals which are worthy of consideration. In the coming discussion, I touch further on some of the issues that are relevant to notions of explanation, and I further claim that some of the proposals considered here succeed in being wrong, but many of these points will not resurface until chapter 5, when I investigate certain of my own claims along these lines.

In this section I deal with certain traditional ‘non-theoretical’ accounts which have been proposed as explanations for lenition; these are at times formulated relatively vaguely which does not facilitate their predictive power. Certain other accounts that I deal with are formulated much more precisely and impressively, and these are accorded commensurately more space for discussion.

The accounts that I deal with are intended to represent a range of approaches which view lenition from several different perspectives. Some of these perspectives are quite ‘phonetic’ in that the account relies on purely articulatory or acoustic factors and others are more ‘phonological’ in that they appeal to more abstract concepts which are not directly observable in articulation or in the signal. It is sometimes difficult to compare approaches of these two types because they proceed from different fundamental assumptions, and are thus only partly comparable. They can all be considered on their own merits and in their own terms, however, and we can evaluate the degree to which they engage with the observations and generalisations which were recognised for lenition processes in section 3.3.2.1. My conclusions on the issues are presented in section 3.3.2.3.

3.2.2.2.1 Lenition scales as explanations

The simple preliminary definition of lenition given in section 1.4 was framed in terms of lenition trajectories, which are also known as ‘lenition scales’. Anything which moved a segment down a lenition trajectory was counted as a lenition process and this provided, at least, a useful way of approaching the concept. I argued above, however, that they can be seen

to conceal as much as they reveal because they do not encourage us to consider the individual lenition processes in their own right. Various authors (for example Escure 1977 and Foley 1977) have disagreed and have proposed that these very lenition trajectories can, or even should be incorporated into phonological theory, proposing that the scales can be seen as explanatory notions and are themselves the fundament of lenition theory.

These approaches are phonologically 'abstract' in the sense that they relate all lenition processes to a notion of *phonological strength* which is not based on any articulatory correlate but which can be observed in the behaviour of phonological segments and which is exemplified by the relationship between segment-types on lenition, or 'strength' scales.⁷³ The concept of strength is indeed an enticing one, for if it can be shown in a principled way that such a notion exists and that all lenition processes represent a reduction in it, then the phonologist can claim achieved success in explaining lenition, to the extent that it will be clear what counts as a lenition process, and why. Lenition processes could be unified as a process-type in a non-arbitrary fashion.

However, approaches such as this typically seem to rely on the approach of (i) noting what happens in historico-synchronic phonology (ie, what kind of processes are innovated), (ii) writing trajectories or scales like that in (1.2) on the basis of this, and then (iii) using these scales to 'explain' further examples of the types of processes involved. It is generally not clear what the notion of phonological strength can be reduced to, and, at least in simple formulations, it seems to be derived from the processes which it is intended to explain. The approach does have the advantage that it can group together analogous processes from a range of languages, and the comparison that this allows can be enlightening, but it seems to be inadmissible as an explanatory lenition theory of 'what lenition is' because of its circularity.⁷⁴

A similar idea to the 'strength scale' notion is found in Vennemann (1983), where Vennemann proposes to accord Popperian scientific status to such scales (or, at least, a list of process types) by stating that, once a 'closed catalogue' of such changes has been devised, this catalogue makes predictions as to what is an impossible change (ie, something that is not on the list). Vennemann's claim here might be true, but it is not clear that this is an interesting

⁷³ Foley is clear on this point "[a] phonological theory based on the phonetic composition of the manifest elements would exhibit the reductionist fallacy and fail to yield insight in to the nature of language " (1977, 25).

⁷⁴ Foley's work went further than this; he considered certain other types of process as well, and seems at times to be making interesting predictions about what can constitute a possible lenition, for example that segments at certain places of articulation will only lenite if segments at other places of articulation also lenite. However, these predictions often seem to melt away when confronted with data, something which has often been done (eg, in Cohen 1971, Smith 1981, Harris 1985, Anderson & Ewen 1987, Harris-Northall 1990 and the extensive survey of Kirchner 1998) and, in replies, Foley has claimed that such predictions were never meant and that his critics "merely misunderstood my theory" (Foley 1981, 601).

prediction because the implications of falsification are not particularly important. All that would need to happen to reassert the list's credentials would be to add a new process to the list. Lenition scales are arguably better than a simple list because they typically only allow movement in one direction and hence are more restrictive, but as we have seen, the whole approach may be more descriptive than explanatory. Vennemann's approach does not tell us anything about why certain changes are impossible (about why they do not belong on the list) and it does not show any kind of commonality among the various kinds of process-types.

It is indeed possible to construct scales of phonological strength from lenition and similar processes, and these often show a reasonably close fit with sonority hierarchies that are derived from the possibilities of segmental phonotactic combinability in syllables, but a problem with scales of this sort is that different theorists devise different scales, and this implies that the facts are not clear enough to allow us to draw up complete scales securely. The place of nasals on strength scales is a matter of some debate, as is the relative strength of 'voiced' stops and 'voiceless' fricatives. It is also clear that such scales have not been formulated to respect laryngeal realism and it is not obvious that they can survive this recognition of the true laryngeal phonology of languages: if it is not clear how fricatives and stops interact with only two series, the prospect of another underlying series will likely complicate matters further. Scales of this sort do not offer any means of connecting with the observations and generalisations of section 3.3.2.1 and while they can be a useful heuristic, their explanatory promise seems limited.

3.2.2.2.2 Is lenition just assimilation?

At times it has been proposed that lenition processes are simply kinds of assimilation processes. If true, this would rob them of any inherent interest in their own right, as there would, strictly speaking, be no such things as lenition processes, only assimilations that we might sometimes refer to as 'lenitions'. This position is in fact perhaps most commonly recognisable in writers who do not really make use of lenition as a concept. For example, Goyvaerts (1975) simply lists 'voicing of intervocalic consonants' and 'loss of plosion in intervocalic consonants' (which are types of sonorising and opening lenition for us) as types of assimilation and does not consider the matter further. A similar approach is taken in Jeffers & Lehiste (1979), Kiparsky (1988) and, as Kirchner (1998) reports, Harris (1984) accounts for the type of Spanish data discussed in section 3.2.1.2 in this way.

This position seems problematic for us here. Assimilations are a straightforward set of processes which involve the spreading of segmental material from adjacent or nearly adjacent segments, and if this is all that lenition is, then there is probably little else left to write on the topic. This is not the obvious conclusion, however. It is not clear that the types of processes that Goyvaerts lists are necessarily assimilations. If spirantisation is to be accounted for in this way, then it must be described as the spreading of a [+continuant] value from the surrounding vowels. There are substantial problems with such an analysis, however. On the model of phonology developed in this thesis, it is inadmissible, of course, as we recognise no such feature [\pm continuant]. The real problem with this analysis is not simply this formal point, however, but is rather due to the reasons why the feature [\pm continuant] is rejected. For a start, it does not seem obviously correct to claim that the degree of stricture involved in fricatives, or even in approximants, is the same as that involved in vowels, so there is no argument here for characterising them both by the same value of a single [\pm continuant] feature. Secondly, it is not clear which value of [\pm continuant] should be associated with sonorants such as nasals and laterals. We have seen that these can have different effects in terms of spirantisation, for example in the Spanish case in 3.2.1.2, where it occurs between a lateral and a vowel in /b, g/, but not in /d/.⁷⁵ The very fact that spirantisation does not just occur between clearly '[+continuant]' segments casts doubt on such an analysis, as does the fact that it is not clear that such processes only occur word-medially if we consider GCS₁ and the processes in Proto-Iranian and Kannada in section 3.2.1.

It also seems clear that assimilation cannot be proposed to account for affrication and debuccalisation because the output of these processes has little in common with the environment in which they occur, and, as we have also noted, it is frequently assumed that the 'final stage of lenition' is deletion and this, too, can hardly be seen as assimilation. If we are right in assuming some kind of connection between all these processes, then we cannot simply equate lenition with assimilation.

3.2.2.2.3 'Ease of articulation'

Perhaps the oldest and most pervasive definition or explanation for lenition processes is that they involve a reduction in articulatory effort. In fact, this has been connected (for example in

⁷⁵ As we saw in section 3.2.1.2, the term 'spirantisation' here may well mean 'approximantisation' in this case, but the argument is the same, as mentioned in this section.

Trask 1996) with the idea that assimilation is involved, on the argument that assimilation between segments means that some of the articulatory activity which would contrast two neighbouring segments is removed. A range of connected notions are often invoked in the discussion of this idea, including ‘relaxation of effort’, ‘laziness’, a ‘principle of least effort’, a ‘law of economy’. In contrast to the more abstract, ‘phonological’ approach which was seen in the strength scales approach, and possibly in the assimilation approach, the reasoning which has typically been brought to bear in connection with the ideas discussed in this section is largely concrete or ‘phonetic’.

Ideas of this kind have a long history as proposals for the explanation of many aspects of linguistic change, including lenition, and have been connected with the work of Paul (eg, 1886) and Jespersen (eg, 1922), as noted by Lass (1980) and Labov (2001), who writes “[t]his principle seems to have been a part of linguistic thinking about change from the very beginning” (2001, 16).⁷⁶

It is indeed apparent in Thurneysen’s (1898) idea of what he was referring to when he invented the term ‘lenition’ as he states that all such processes seem to him to involve a “reduction in articulatory intensity” (1898, 43).⁷⁷ It may well be impossible to trace the originator of this idea. Caudmont (2001) claims that Martinet (eg, 1955) replaced an old ‘law of least effort’ which he attributes to Passy (1890), with a concept of ‘linguistic economy’. Passy (1890) does indeed consider the idea that there is a “general principle in phonetic tendencies” for effects to be due to a “law of least effort” (1890, 227).⁷⁸ Passy’s point, however, is that this had previously often been connected with a ‘principle of economy’ which functions to remove what is unnecessary from language (for example, to remove effort, where it is not necessary to preserve meaning). Belying the idea that older work is oversimplistic, he writes that this is a complex notion and should not be viewed as explanatory in a simple way.

These ideas go further back, however, as does the discussion of how insightful they are: Passy (1890) refers back with deference to Sweet (1888), who writes that “[t]here seems, indeed, reason to doubt whether the inherent ease of articulation has much to do even with

⁷⁶ Trask (2000) attributes the introduction into linguistic theory of the idea that a *Bequemlichkeitstrieb* (‘drive for comfort’ or perhaps ‘pressure for ease’) can motivate linguistic change to Georg von der Gabelentz in the nineteenth century.

⁷⁷ In the original, he writes that “mir in a l l e n Fällen die Minderung der Artikulationsintensität die Grundlage zu bilden scheint.”

⁷⁸ In the original, one of Passy’s “[p]rincipes généraux des tendances phonétiques” is that “[l]e langage tend constamment à se débarrasser de ce qui est superflu”; he continues that this principle “est ordinairement considéré en même temps que la recherche des articulations « faciles », et réuni avec elle sous le nom de *loi du moindre effort*” (1890, 225-227).

isolative change. As a general rule, all familiar sounds seems easy, all foreign ones difficult and harsh” (Sweet 1888, 49).

It is clearly not the case that the whole of lenition is explained by simply equating it with a reduction in articulatory effort. There is some intuitive appeal to the idea for at least some of the processes that we have identified as ‘lenition processes’, however. For example, if debuccalisation involves the loss of oral articulation (following Lass 1976 as discussed in section 3.1.3.1.8) then this would indeed presumably require less articulatory effort on the speaker’s part. However, it is not clear that affricates require less articulation than stops and in fact, it is also not clear that spirantisation involves a lessening of articulatory effort, irrespective of whether the output is a fricative or an approximant. This issue was discussed on the Optimality email discussion list⁷⁹ in 1997, and two contradictory positions were proposed for the scenario of intervocalic spirantisation of stops: (i) that it requires less effort if the tongue does not have to raise all the way to the palate (or if the lips do not have to close completely) because two flanking vowels require the articulators to be in open approximation, so if less approximation is introduced into in the stop, this will lessen the effort required to produce it and a fricative or approximant will result. Opposing this is the proposal that (ii) simple ballistic articulations do not require much effort and that it is less effortful to allow an articulator (tongue or lips) to pass through the whole of the articulatory space and hit their opposing passive articulator, which then cannot miss if they travel all the way (because of the physiology of the mouth) than it is to pull short of this and engage in a period of relatively close approximation.

The problem here is that both arguments rely on notions of ease in articulation, but that they come to opposite results and the status of spirantisation as a case of reduction in articulatory effort is therefore in doubt. It partly can be seen to rely on which metaphor a theorist prefers: either it is easier for speakers to aim for a target and miss (which would mean, with exemplification for velars, that VxV is easier than VkV) or that it is easier for speakers to simply launch out for target (which you can’t miss) and not worry about stopping before you get there (which would mean that VkV is easier than VxV).

There are further problems with the notion. A process such as $p^{\circ} > b$, described in the history of Spanish in (3.38), actually involves the addition of articulation. If /b, d, g/ are specified for |voice| themselves, then the process which leads to a diachronic change in the

⁷⁹ For detail of the list, see <http://rucss.rutgers.edu/ROA/o-list.html>.

underlying form of the segments has actually increased the articulation that is involved in the production of the segment by itself.

Given the above issues, we may well question whether all lenitions are simply explained by claiming that they are due to 'ease of articulation.' If we had a clearly articulated theory of what 'effort' means and how it can be measured, then serious work with the idea could be undertaken, but even then it does not seem that all of the lenition processes that we have dealt with would count as a reduction in it.

There has been work which has sought to provide a precise definition of how we could really understand the notion of what 'reduction in effort' might mean precisely and it is worth addressing these, as they seek, at times with impressive results, to provide a proper definition for an idea that phoneticians and phonologists have long wanted to use in their approaches to explanation. I discuss such work in section 3.2.2.2.3.1.

One attempt tries to reduce the reduction of effort in lenition processes to a simple primitive, which would qualify the idea as being explanatory, but it is not clear that the attempt is successful as it remains rather vague. In his proposal, Bauer (1988) seeks to relate all lenition processes (along with vocalic processes such as $a > \text{ə}$) to reduction in the activity of the posterior cricoarytenoid muscle, but few specifics are adduced to illustrate that all the stages along a lenition trajectory involve reduction in the activity of this muscle. In fact, Bauer (1988) points out that there is not any published data which backs up the putative correlation between the activity of this muscle and most lenition processes, and also admits that voiced fricatives involve more activity in the muscle than do voiced stops. Such a reductionist approach is admirable in principle, but it seems to fail on the specifics. One recent proposal has sought to be very specific in this regard and has produced some impressive results. I turn to this approach to 'effort' in the next specialised subsection.

3.2.2.2.3.1 Kirchner (1998, 2000)

Kirchner (1998, 2000) has sought to provide a robust definition of articulatory effort and to use it to provide explanatory accounts of certain lenition phenomena in a substantial contribution to lenition theory. Kirchner's work is based on a set of generalisations about the behaviour of certain types of segments in lenition which are partly based on a survey of languages for cases of lenition in Lavoie (1996) and partly on his own. It was argued at the start of section 3.2.1 that surveys with such breadth (Kirchner includes data from 272

languages) need to be supplemented by studies which consider the lenition phenomena in depth, but this work is clearly important. We return to the implication of some of his generalisations later, but we can note here that much of Kirchner's work focuses on the behaviour of geminates in lenition (this is particularly the case in Kirchner 2000). We have only briefly encountered geminates in this thesis, as a product of the HGCS₁, and we will not change focus here. We return to some of his generalisations in the next major section (3.2.3) when we consider the phonological environment in which segments can be affected by lenition (it is argued there that the effects observed in geminacy are part of a wider generalisation) and we turn to the specific case of the HGCS₁ in chapter 4. In this section, I chiefly examine some of the interpretations and conclusions that Kirchner proposes for some of these generalisations.


Kirchner's approach to lenition is based on the notion that lenition processes reduce the articulatory effort that speakers need to expend in order to pronounce sequences, and he considers a range of physiological and physical evidence. We cannot discuss all the details here, but briefly, his understanding of effort relies on the amount of adenosine triphosphate ('ATP', which he describes informally as 'muscle fuel') that muscles have to consume in order to execute any particular articulation. Kirchner claims that information regarding the amount that is consumed in any articulation is fed-back to the nervous system to the extent that speakers can develop a knowledge of the amount of ATP which any articulation requires. This knowledge, he claims, is the basis of psychologically real phonological analysis as to which combinations of segments are effortful and which are less effortful.


One of the generalisations that he recognises is that "[n]o process converts a voiceless segment (geminate or otherwise) to a voiced geminate obstruent" (1998, 126 and 2000, 510). Kirchner then goes on to illustrate how vocal cord activity during complete oral closure is quite effortful for physiological reasons and would thus involve the consumption of more ATP than the lack of such activity, hence, for example, /gg/ is more effortful than /kk/. Kirchner extends this approach to account for other aspects of the behaviour of geminates and certain aspects of spirantisation, but he does not deal with the interaction of affrication and spirantisation, nor does he consider the richer understanding of laryngeal specifications which has been developed in this thesis.

The approach to articulatory effort developed by Kirchner may well provide a motivation for the diachronic innovation of certain types of lenition process, and he makes serious attempts to be specific as to how effort can be measured, but the way in which Kirchner

formalises his insights is far from minimalist. He assumes an Optimality Theoretic model of phonology and imputes to it the ability to evaluate the ATP consumption that any phonological configuration will require. This is built into a “scalar effort minimization constraint, LAZY, (which generally favors reduction of articulatory gestures, ideally to Ø)” (Kirchner 2000, 529). This constraint knows the potential ATP consumption for every phonological configuration, and it weighs up all candidate outputs in terms of this and ranks then by giving those which require more effort more asterisks than those which require less effort. As is natural, given the OT nature of the analysis, this constraint needs to be balanced with faithfulness constraints, and, following McCarthy & Prince (1995), Kirchner (2000) posits both general and positional IDENT constraints for every feature and when these are ranked above LAZY, no lenition occurs, as shown in the first tableau in (3.55). When LAZY is promoted above these faithfulness constraints, lenition occurs, as shown in the second tableau below:

(3.55)

/b/	IDENT (cont)	LAZY
 b		
β	*!	

/b/	LAZY	IDENT (cont)
b	**!	
 β	*	

One problem with this analysis is the large number of constraints which are postulated. As well as the all-important LAZY, Kirchner uses general IDENT constraints of the type shown here, and also ‘positional IDENT constraints’ such as IDENT (cont/onset), which are needed to account for lenitions which are sensitive to phonological environment. Also, because of the OT notion of ‘Richness of the Base’ (which was briefly discussed here in section 3.1.2), Kirchner also requires ‘fortition’ constraints, such as (*[+cont, -son]/#__), which force multiple possible inputs to surface as the correct output. The analysis is made less compelling

by the large amount of subsidiary machinery which it requires, much of which seems to restate the observations of the analyst (eg, ‘stops do not spirantise at the start of words’).

For Kirchner, the analysis of any actual lenition can involve a substantial amount of theoretical machinery and the innovation of any process requires this machinery to be reranked. As we saw at the start of section 3.3.2.2, this reranking is in no way explanatory, because in terms of the grammatical model, it is just as likely to be demoted on the hierarchy as it is to be promoted.

A further problem is the notion of LAZY itself. The computational power that is attributed to this constraint is immense, and it is not clear how the mechanism works through which the constraint receives feedback from articulators when they carry out the articulations that it is to forbid and then proceeds to calculate the amount of ATP which would be required by them if a speaker were to produce the effortful articulations. For Kirchner the whole of lenition is placed in the LAZY constraint and the asterisks that it produces as violation marks, but the constraint is not like most OT constraints, which simply militate against a particular output form, and this makes it a little dubious.

Kirchner’s work seeks to make the thinking behind the ‘ease of articulation’ argument measurable, but it is not clear that the formalism into which it is placed adds much to the basic idea that (i) there are pressures of articulation and (ii) there are pressures which militate against their affecting phonology, which has always been a part of the ‘ease of articulation’ position. While the measurement of ATP-use, or its OT formalisation provides a new way of conceiving of the ease of articulation, the real argumentation behind the position lies in showing that lenition outputs involve less effort than lenition inputs and this is still not always clearly the case. It seems that the model provides no reason why certain segments are lenited into others and makes no predictions as to what is a possible lenition process because no predictions can be made about what is a possible constraint reranking. The approach is also incapable of recognising any of the observations that were identified in section 3.3.2.1.

We return to certain aspects of Kirchner’s work below, but we leave the ease of articulation argument here. Discussion of the ideas involved is often unsatisfactory, because it can often be vague and, while Kirchner has produced a formalised account, the precise way in which a pressure to produce more ‘easy’ articulations can be seen to cause lenition is somewhat unclear in the ideas discussed in this section because the pressures are all simply placed in the constraint LAZY. The problems identified with the ‘ease of articulation’ argument generally above still seem to hold.

The notion that lenition involves a reduction in effort has yet to provide a full account for what occurs in lenition processes and it does not seem to have shed as much light on ‘what lenition is’ as we might hope for, in part, at least, because it misses certain generalisations. In the next sections, I turn to some accounts which seek to model the processes explicitly in an attempt to understand the precise ways in which segments are affected in lenition and to recognise an underlying unity which makes them all examples of lenition. The discussion there will be of a quite different form to that in this section because the burden of explanation (to the extent that anything in historical phonology merits the term) is placed on segmental representation. This will connect more clearly with the concerns of section 3.1.3, where it was argued that a clear understanding of phonological structure is vital to the understanding of phonological processes.

3.2.2.2.4 Articulatory Phonology

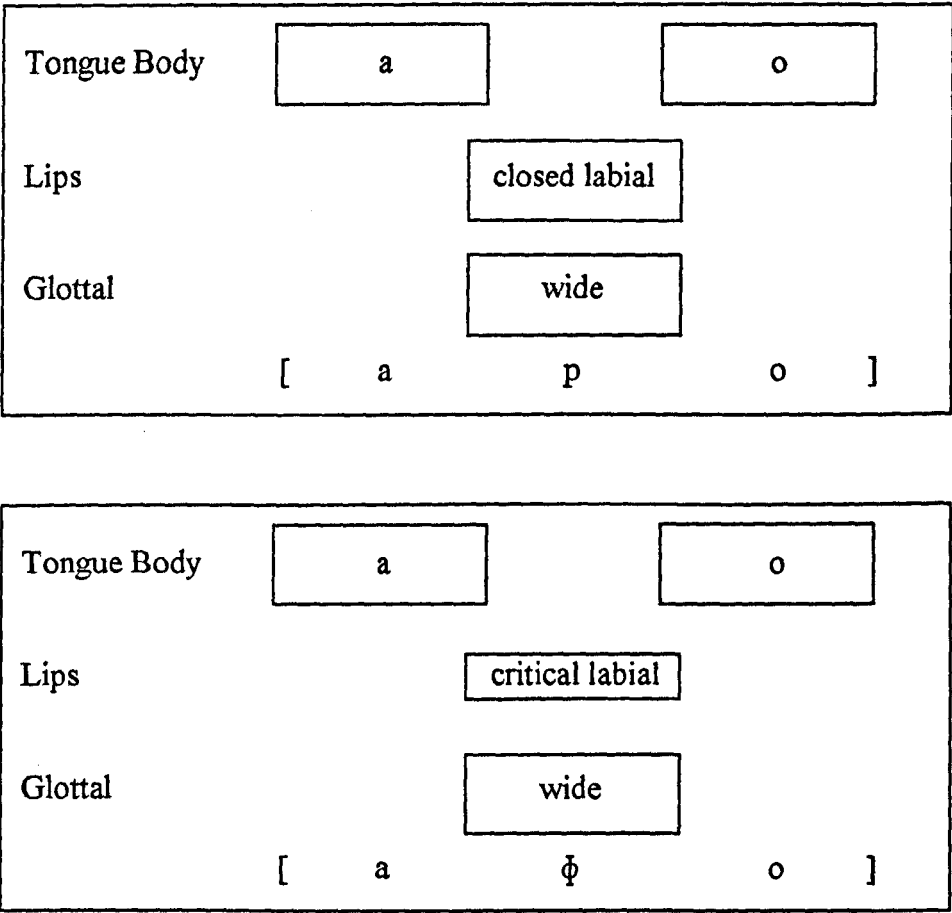
In section 3.1.3, we very briefly encountered some of the ideas of the Articulatory Phonology framework. The point there was that the articulatory gestures employed in the model have a certain amount in common with the elements which I argued are the best basis for phonological structure. The gestures used in the Articulatory Phonology model are certainly closer in principle to elements than the features which Kirchner employs, along with many others, as we saw in the phonological rules of King (1969) and Lehmann (1973) and the constraints of Jacobs (1994) which were discussed at the start of the current section (3.3.2.2). Subsegmental representation plays an important role in Articulatory Phonology, so it is possible that the model can contribute to our understanding of lenition to the extent that the other approaches discussed in this section so far have not been really able to, by allowing us to model the processes in an insightful way.

The elements which I adopted in section 3.1.3, principally from the Government Phonology and Dependency Phonology traditions, are typically conceived of as wholes which cannot be divided or diminished. This is one area where there is a clear contrast with the approach of Articulatory Phonology. The originators of the model (Browman & Goldstein 1986, 1992) worked with very restrictive ideas, so that the deletion of a gesture, for example, was not possible, unlike in Government and Dependency phonologies, where elements (or components) can be lost from the structure of a segment. They do, however, allow for the diminution of gestures. In fact, this is one of the key mechanisms that are available in the framework for the modelling of phonological processes (the other is for gestures to overlap,

which can be used to model assimilations, for example, but also for apparent deletions as one gesture can be thought to ‘hide’ another).

McMahon (2000a) shows how this notion of gesture-reduction can be used to model spirantisation by comparing the ‘gestural scores’ (which are used in Articulatory Phonology as phonological representations) of [apo] with [a ϕ o]. The scores are reproduced as (3.56) below:

(3.56)



We can see from these scores that /p/ is modelled as consisting of a ‘closed labial’ gesture and a ‘wide glottal’ gesture. These do not map onto the elements in an entirely straightforward fashion, although ‘wide glottis’ can be equated with |spread|. In terms of the elements proposed in section 3.1.3, the gestures ‘closed labial’ and ‘critical labial’ would need to be split into two parts: both could be seen to consist of |labial|; the key difference comes in the modelling of manner. It is in this ‘manner’ gesture/element that the lenition is modelled, as a physical reduction in magnitude.

The way in which Articulatory Phonology thus allows the lenition of stops to fricatives (or approximants?) to be modelled as a reduction in the ‘manner’ or ‘stricture’ of a gesture can be seen as insightful; it could become truly interesting as a lenition theory if it could be extended to other lenitions - if all lenition processes can be modelled in this way, then the model could claim to provide an insight into lenition as a whole. This would be a much more overt unification than Kirchner’s use of interactions of faithfulness constraints with LAZY, which might be seen to hide as much as it reveals.

However, this does not seem to be possible. Debuccalisation, for example, can hardly be seen as simply the reduction of a gesture, rather, it would seem to involve the deletion of the tongue/lips articulatory gesture. Strict Articulatory Phonology does not allow the deletion of gestures, however, although it is possible that a more ‘phonologised’ model which still works with gestures, such as that in McMahon, Foulkes & Tollfree (1994) might be able to work with such an idea, but then the unification is lost. Hind (1996), who is working in a slightly phonologised Articulatory Phonology seems unable to make this step, however, and analyses debuccalisation of /s/ to [h] in Sanskrit in an entirely different way (where ‘TT’ stands for ‘tongue tip’): “[d]ebuccalisation means that /s/ has an alternative form which lacks a TT gesture” (1996, 213). This accepts the Lassian analysis of the glottal fricative as simply glottal activity (as a ‘wide glottal’ gesture), but instead of viewing debuccalisation simply as a dynamic debuccalising process, Hind (1996) seems to be suggesting that /s/ and /h/ are entirely separate alternating segments, which rather misses the naturalness in the derivability of one from the other.

It seems rather unclear whether a coherent ‘lenition theory’ could be worked out in the Articulatory Phonology approach. We may well wonder whether this is a problem. If we need to formally unify lenition processes as one type of action in a phonological model, then Articulatory Phonology seems to fail. I discuss below how serious a failing that is. To the extent that we have discussed it here, however, it seems that, in the strict approach, the framework cannot even model all the processes which we have dealt with in this thesis, and this surely is a problem. Although these could be remedied through further ‘phonologisation’ of the type discussed above, this will likely make the model of segmental structure and processes more like those of Government and Dependency Phonology, to which we turn now. We will see that, as was explained in section 3.1.3, the model that I have adopted, while inspired by them, has clear differences from both models, in terms of both theory and practice, especially with ‘strict’ interpretations of the frameworks. We shall see that explicitly

articulated 'lenition theories' have been proposed in both Dependency and Government phonologies, although the proposals are quite different in their details, and conflict in quite fundamental assumptions.

3.2.2.2.5 Dependency Phonology

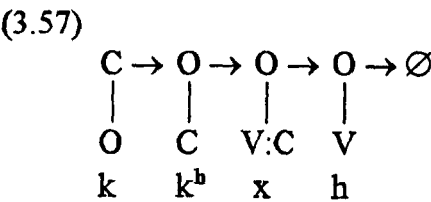
More space was given in section 3.1.3 to ideas about segmental structure from the traditions of Dependency Phonology than was given to those of Articulatory phonology, but in fact, the discussion there only dealt with ideas about one of the aspects of subsegmental structure which are usually recognised in the theory: those of the 'articulatory gesture'. The understanding of 'gesture' here is not really comparable with the gestures of Articulatory Phonology, as just discussed. The gestures of Dependency Phonology give structure to the segment, grouping the elements (or 'components' as they are more normally called) into units which correspond very roughly to traditional notions of 'place' and 'manner' and the like.

Lass (1984a) recognises three gestures: categorial, articulatory and initiatory. Anderson & Ewen (1987) recognise similar distinctions, but with more structure. They have two major gestures 'categorial' and 'articulatory', which are each split into two: the categorial gesture is comprised of the 'phonatory' and 'initiatory' sub-gestures and the 'articulatory' gesture comprises the 'locational' and the 'oro-nasal' sub-gestures. An explicit understanding of lenition has been developed which principally refers to the units in the categorial gesture, which collates factors in a way that Lass (1984a) summarises thus "[t]he categorial gesture defines (roughly) the degree of 'consonantality' or 'vocalicity' of a segment" (1984, 282). The fact that Anderson & Ewen have split this into two will be relevant below. The initiatory (sub-)gesture deals with laryngeal specifications and the articulatory gesture deals with place and nasality. The gestures can be treated independently from one another and it is natural here to focus on models of the categorial gesture because that is most closely linked to what happens in lenition.

It is proposed that the phonatory categorial gesture determines relative consonanticity through the use of two units, $|V|$ and $|C|$, which by themselves represent a vowel and a voiceless stop, respectively. These two combine in various ways with several types of dependency relations (which can be seen as extending the notion of 'head' which was used in section 3.1.3). A voiced stop combines $|V|$ and $|C|$, with $|V|$ as a dependent and $|C|$ as a 'governor' (the representation in Anderson & Ewen 1987 is $|C=V|$). A voiceless fricative combines $|V|$ and $|C|$ with a different dependency relation (both are dependent on each other,

giving a representation $|V:C|$). Voiced fricatives combine $|V:C|$ with a further occurrence of dependent $|V|$, giving $|V:C=V|$. We do not need to deal with the details of the representation of all segment types here, but the point is that the prominence of the two units ($|C|$ and $|V|$) vary in the characterisations and this allows a characterisation of many cases of lenition as an increase in the prominence of $|V|$. The ‘start’ of most lenition trajectories is the voiceless stop, ie, $|C|$, voicing of this stop involves the addition of one $|V|$, spirantisation involves the addition of another (in a different dependency relation). It is proposed that a further stage of lenition changes a fricative to an approximant and this is characterised by the promotion of $|V|$ in the structure of the segment (in terms of dependency relations). This approach moves towards the unification of lenition as a phenomenon by describing lenition processes as involving the same ‘kind’ of change in the structure of a segment and it thus holds out the promise of producing a platonic lenition theory.

But in fact, Lass (1984) has two separate lenition trajectories (and the approach in Lass & Anderson (1987) seems the same). The problem has to do with laryngeal specifications but this is not even simply the fact that laryngeal realism is ignored. The approach splits up aspects of what we treated in section 3.1.3.1 as a unitary laryngeal category into the parts of the ‘categorical’ gesture just described and aspects of the ‘initiator’ gesture, which are chiefly characterised through the use of a further basic unit: $|O|$, which is glossed as ‘glottal opening’. This in fact forms part of all segments (apart from glottal stop) but it appears with differing dependency relations. The key result of this is that the difference between an unaspirated voiceless stop and an aspirated voiceless stop is represented by a difference in dependency between $|O|$ and $|C|$. Lass’s (1984) lenition trajectory which involves $|O|$ is shown in (3.57), with his exemplification for velars, where the vertical difference indicates a difference in dependency (the top line indicates governors and the bottom line shows dependents):

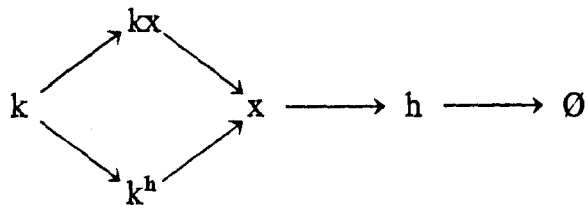


This trajectory shows a different picture to the notion of ‘lenition as promotion of $|V|$ ’, although that idea still plays a role for $k^h > x > h$. The first stage seems peculiar in that it shows a neutral stop as ‘stronger’ than an aspirated stop when, as we have seen in section 3.1.3.1, this actually involves the addition of a laryngeal spreading gesture, whereas a neutral

stop has no glottal activity of its own. In fact, it could be argued on various grounds that /k^h/ and /k/ should be the other way round of a lenition trajectory.⁸⁰ In any case, the $k \rightleftharpoons k^h$ (or vice versa) is different to other lenitions, for example, $x \rightleftharpoons h$, as it involves a change in the prominence of |O|, not of |V|. Therefore the unity of lenition is lost.

Aspiration, thus indicated, has an odd relationship to affrication. The tradition in such work is to include affrication as a case of lenition. Lass (1984) writes that in terms of the stages of lenition between stop and fricative, affrication is “an ‘alternate’ route to aspiration” (1984, 292), and as a full picture of the relationships between stops, affricates, aspirates and fricatives, Lass (1984) proposes the trajectory reproduced (in a slightly adapted form) in (3.58):

(3.58)



(exemplified for velars from Lass 1984, 178)

Affrication is clearly characterisable on Dependency Phonology assumptions as an addition of |V|, but the picture in (3.58) clearly misses the connection between aspiration and affrication which was recognised in section 3.3.2.1. Other aspects of lenition, on a conventional understanding, are quite well captured on this picture, but the prospect of unification does not seem to have been achieved. A further example of this is that debuccalisation to [h] involves the loss of the articulatory gesture as well as the loss of |C| in the categorial gesture which is shown in (3.57). Given this, we may question the need for the complex representations of manner which are adopted in the theory, especially as they are motivated, at least in part, by the patterns of lenition. The representations adopted in section 3.1.3, motivated by those used in Government Phonology, are more minimalist and, as we will see, are no worse at capturing the patterning of lenition processes than those explored here (and both are better than standard feature theory).

⁸⁰ For example: if such trajectories indicate a lessening of effort then it surely involves more articulation to produce a laryngeal gesture than not to produce one; if they indicate what happens in ‘lenition sites’ then the underlyingly |spread| stops of English deaspirate in such environments (eg, environment C1 [\acute{v} __ (v)] ‘foot-internal’ and environment B [__ c] ‘coda’); if they indicate decrease in salience, then the noise associated with the aspirated release of /k^h/ is greater (ie, more salient) than it is for the release of /k/.

3.2.2.2.6 Government Phonology

The proposals which most influenced the model of segmental structure which was adopted in section 3.1.3 were made from within the theory of Government Phonology, as was made explicit in that section. The wider theory is founded on proposals which are not discussed here and are not necessarily adopted (see, for example, Kaye, Lowenstamm & Vergnaud 1985, 1990 and Charette 1991 for details). They focus on a set of licensing relations which are proposed to exist between segments (or between prosodic skeletal positions). These dictate the notions of possible phonological domain in the framework and are also made accountable for phonological processes; they are called 'government'. The elemental model was designed to work with phonological government to provide a full picture of phonology and was elaborated in Kaye, Lowenstamm & Vergnaud (1989, 1990), Harris (1990, 1994) and other work. While many of the basic tenets of the elemental approach are shared with other models, as we have seen, the way in which elements are used in consonants, especially, is a novel development. Working in this tradition, Harris (1990, 1994, 1997) has developed a detailed lenition theory. I examine certain key aspects of this in the coming subsection.

3.2.2.2.6.1 Harris (1990, 1994, 1997)

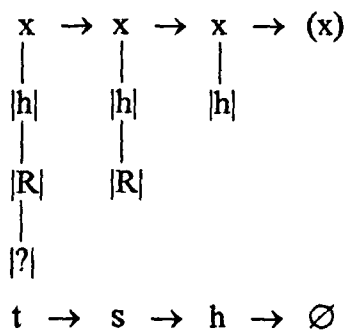
There are two aspects to Harris's approach and, although they are connected, they are also disassociable and I deal with them separately here. In this section, I focus on the aspects which are relevant to the description of the processes involved. I turn to the proposals made to account for the interaction between the processes and the phonological environment in which they occur in section 3.2.3. The first exposition of this approach to the facts on lenition was Harris (1990), since then the ideas have been developed in Harris (1994 and 1997), further work in this tradition includes Brockhaus (1995a) and Bloch-Rozmej (1995). In this discussion, I focus on the presentation in Harris (1994). The approach seeks to provide a unified characterisation of lenition processes and, given the nature of the theoretical framework in which it is proposed, it is an abstract, phonological approach which does not seek to connect with phonetic notions such as ease of articulation, and like the Dependency Phonology approach just described, it does not rely on stipulative, circular strength scales.

The basic approach relies on notions about the elemental composition of segments which are similar to those proposed in section 3.1.3, although there are certain key differences. The

elements which Harris assumes are symbolised by single letters or symbols and are not written within vertical slashes, but I retain the slashes from section 3.1.3 for typographic clarity. We have already seen in some detail in section 3.1.3.1 that Harris assumes the laryngeal elements [H] and [L], which I transcribe as [spread] and [voice]. In terms of ‘place’ elements, he assumes that some of these are the same elements as those found in vowels: he proposes that [labiality] is the element [U] (which by itself is /u/ or /w/), that [coronality] is [R] (a purely consonantal element which by itself is /r/), and that [dorsality] is [ʔ] (which by itself is /ə/ or /ʉ/).

The key differences are in terms of the ‘manner’ elements, although [nasality] is straightforwardly [N]. Harris uses [ʔ] for [occlusion] and the closest equivalent in his system to [frication] is [h]. We saw in section 3.1.3.1.8 that the facts of fusion show that it is the element [spread] which is interpreted by itself as the segment /h/, but Harris assumes that /h/ consists of [h] (our [frication]). This assumption, along with the other factors associated with elementhood (which I have taken over in this thesis) allow Harris (1994, 124) to characterise a relatively full lenition trajectory in the manner shown in (3.59):

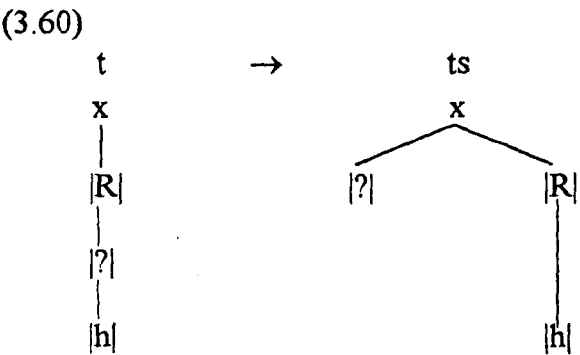
(3.59)



This assumption is possible because Harris reinterprets [h] (which was originally proposed by Kaye, Lowenstamm & Vergnaud, as were the other elements) as ‘noise’ which is found in the frication of fricatives and affricates, but also, he proposes that it accounts for the release of plosives, hence it is in the structure of /t/. Harris’s analysis of lenition thus builds on the notion of the loss of segmental material, which was introduced to account for debuccalisation in Lass (1976) and which we accepted for debuccalisation in section 3.1.3.1.8. This idea seems like a major step towards the formal unification of lenition processes, and, if the analysis holds good, it is a substantial contribution to lenition theory and would be good evidence in favour of the basic theoretical assumptions involved. The unificatory idea is that

lenition is formalisable as element loss⁸¹ and it has been extended to account for other phenomena, such as tapping (briefly mentioned here in section 3.2.1.1) and glottaling.

However, there are problems with the proposal. In terms of the processes discussed in this thesis, affrication, for example, cannot be captured as the loss of elements. Harris (1990) analyses affrication as in (3.60), to produce the type of contour segment which was adopted in section 3.1:



This is formally unlike the others kinds of lenition processes, in that it does not involve the loss of elements. Harris (1994) tries to overcome this problem by claiming that affrication has in common with other types of lenition a reduction in the degree of fusion contained in the segment, but this is clearly a retreat from the strong and appealing position that lenition equals element loss, and it is unclear why other elements are never ‘de-fused’ in this way. In fact, the ‘breaking’ (as Harris informally names it) shown in (3.60) is rather similar to the way in which diphthongisation processes can be modelled in the theory, but these are not viewed as lenition phenomena. The facts of voicing also present a problem for Harris, and we investigate this in the next section (along with certain other aspects of lenition discussed).

There are also certain theoretical problems with Harris’s proposals. It was argued above that Harris’s use of [h] is problematic. One clear reason for this is that we saw in section 3.1.3.1.8 that /h/ can be fused with plain stops to produce aspirated stops. This speaks substantially against Harris’s idea for the interpretation of the element as ‘noise/frication’, rather than [spread], unless it is proposed that both elements can be interpreted as /h/, which seems methodologically dubious. We might also question the analysis which gives /h/ a role

⁸¹ My use of the word ‘loss’ here may not be strictly in line with the truth of the proposal. In Harris (1997), for example, there is a clear indication that Harris views the elements as still being part of the phonological make-up of a segment, but as being unparsed. For our purposes here, the difference can be seen as terminological.

in stops. As explained above, for Harris, this represents the noise burst of the release of a stop, but it is not clear that any element is needed to provide this. The activity involved in the production of a stop (ie, occlusion) unavoidably entails release if we are to use the oral channel for speech, and any kind of stop release will produce noise if air is flowing through the mouth at the time. As we saw in section 3.1.3.1, a specification such as *|spread|* will give a particularly noisy release as the glottal gesture involved leads to a period of turbulence after the release of occlusion and before voicing can start in a following voiced segment. This all seems to indicate that Harris's proposal for *|h|* cannot stand.

This is problematic for Harris's analysis, because it removes the ability to characterise spirantisation. As we saw in section 3.2.2.1, this may not be so problematic for certain cases of lenitions such as the Spanish case of section 3.2.1.2, which typically involve approximants, and not fricatives, but does seem to be problematic for other cases, such as */p^h, t^h, k^h/ ⇔ /f^h, θ^h, x^h/*. We return to the implications of this in the next section.

While they are perhaps the best attempt that we have seen of the formulation of a theory of lenition, Harris's proposals nonetheless seem flawed. The idea of lenition as the loss of elements is intuitively appealing and seems right for at least certain lenition processes. Other work from different theoretical traditions has made use of the same idea, as we have already seen for debuccalisation; a further example is Grijzenhout's (1995) analysis of Irish initial mutations as the deletion of manner information ('aperture positions').

Like the other approaches discussed above, however, the idea does not seem able to account for all lenition processes. Also like most of the others, it does not connect with the generalisations made in section 3.2.2.1, nor indeed with the observations that derive from the recognition of laryngeal clarity, even though Harris is aware of the issues. This downbeat note concludes our survey of previous approaches to the processes involved in obstruent lenition.

3.2.2.3 If these are lenitions, then what is lenition?

In the last section we first set up a number of criteria that an overarching 'lenition theory' should ideally fulfil and we then investigated a wide range of proposals in the light of them. The proposals had all been put forward as attempts to make sense of the notion of lenition and had the aim of providing at least a unifying definition of what it means for a process to be a 'lenition process'. Nonetheless, we saw that none of the proposals could really claim to

achieve the latter aim,⁸² let alone engage with the former aims, or provide any predictions which might make them potentially explanatory. Phonologists seem to perceive that there is something that lenition processes have in common but they have singularly not succeeded in coming to terms with what that might be, despite a range of ‘concrete’ phonetic and ‘abstract’ phonological approaches.

It seems that we have reached an impasse in our attempt to say something coherent about lenition and I take this as a sign for a change in direction. In this section I take a step back from the attempt to see commonalities and consider each process in its own right. As we will see, this is the approach that has been lying in wait since the phonological representations were proposed for segments in section 3.1. The set of segmental representations which was recognised there has helped to inform the discussion of lenition processes since they were recognised, but we have not investigated whether they can be exploited more fully to provide any insights into lenition. I carry out this task in this section and we will see that they do indeed shed light on what can happen in lenition. I consider each process individually and explore what kind of mechanism could link the representations. The approach that we will recognise in this section will be most like that of Harris, but the considerations discussed here will require us to view certain points in a rather different way to that explored in section 3.2.2.2.6.1. I consider the processes in line with the remarks on process innovation in chapter 1 and on minimalism in phonology in section 3.1, as these conspire to predict what is a possible phonological process. The basic idea here, as we saw previously, is that ‘naturalness’ constrains process inhibition and minimalism maintains that any further process which is abstracted away from such naturalness is not opaque or overcomplicated. I consider the patterning that can be seen in each of the processes in turn and consider whether this allows for any predictions as to what could or could not occur in terms of the processes. Concerns of laryngeal realism will be particularly relevant here. Once each of the processes has been considered individually, we will be in a position to return to the processes as a group to consider whether there is something which might allow us to group them together, after all. We will see that there is, but that it requires a rather different approach to that adopted in the work that we considered in section 3.2.2.2.

In section 3.2.2.3.1, I focus on affrication and consider whether the generalisations observed for this process can be treated insightfully; this will also involve a consideration of certain aspects of spirantisation. In 3.2.2.3.2, I consider further aspects of spirantisation

⁸² Apart, perhaps from the ‘scalar’ approach which essentially defines lenition by saying ‘lenition is lenition’; this is hardly a satisfying definition and doesn’t shed much light on the notion.

which are qualitatively different to those treated in the previous section. In 3.2.2.3.3, I return to the notion of debuccalisation. In 3.2.2.3.4, I discuss ‘voicing’ and related concerns. In section 3.2.2.3.5, I consider all the processes together, in the search for the essence of lenition. In what follows, I exemplify relevant processes at the velar place of articulation unless there is reason to do otherwise.

3.2.2.3.1 Affrication and |spread| spirantisation

We saw in section 3.2.2.1 that there is a correlation between affrication of the HGCS₁ and Liverpool type and the presence of |spread| in the input. This gives us a possible process $k^h \Rightarrow kx^h$. We saw further that there can be a correlation between this type of affrication and spirantisation in segments which involve |spread|. This is shown partly in the fact that in all cases that we have seen at least some of the input segments become |spread| fricatives in the output. It is not immediately obvious from the data how to interpret this, but the Liverpool English data gives us the key. The specialness of this data lies in the fact that it involves a synchronic, observable variable process. We saw in section 3.2.1.8 that the lenitions in Liverpool English give us evidence for a ‘synchronic lenition trajectory’ of the type $k^h \Rightarrow kx^h \Rightarrow x^h$ because in certain environments, all three segment-types are possible, whereas in others, only the stop and affricate are possible and in still others, only the stop. The uniformitarian principle, which was discussed in section 1.2 implies that it is advisable to reconstruct languages in line with what we know to be possible and impossible in contemporary languages. The principle that we should assume minimal quanta, discussed in section 1.2.2.3 implies that it is advisable to reconstruct a series of stages in all languages where there is evidence for it in some languages.⁸³

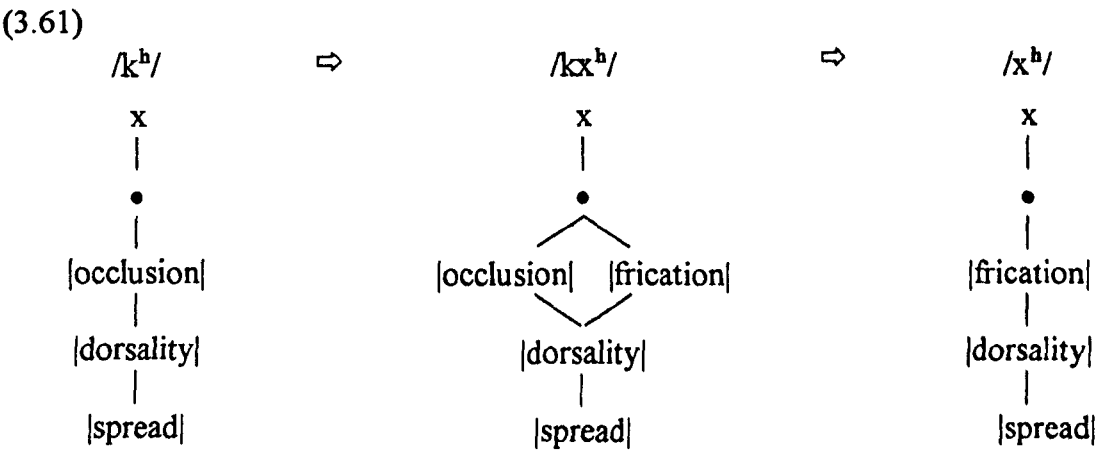
It thus seems that we are on firm, theoretically valid ground to postulate that the expected scenario in such cases is indeed $k^h \Rightarrow kx^h \Rightarrow x^h$. This is not all, however. One of Kirchner’s (1998) generalisations from his survey of attested lenitions⁸⁴ is that they “support a generalisation concerning spirantization: unaffricated stops never lenite to strident fricatives, such as [s] or [f]. Rather, stops typically spirantize to weakly fricated or approximant

⁸³ This is a case of ‘principle overriding data’ see Lass (1997, 219).

⁸⁴ As we saw in section 3.2.2.2.3.1, the generalisations are based on a survey of lenition systems in 272 languages. It was argued above that such surveys must be supplemented with detailed studies of the individual cases. This is certainly true if we want to understand the intricate aspects of patterning, such as those which will be investigated in section 3.2.3, but such surveys can provide us with a basis for generalising over process types.

continuants such as [β, ð, ɣ] or [β, ð, ɰ]" (1998, 99). This ties in with the similar generalisation that Lavoie (2000, 2001) has made for |voice| lenition, as reported in section 3.2.1.2, but can help add to our knowledge of |spread| lenition. This seems to indicate that ‘voiceless’ fricatives⁸⁵ can only be lenition outputs, as in Proto-Iranian, Celtic and most likely Kannada where there is a preceding affricate stage. This would tie in with the observation which was made in section 3.2.2.1 that there are at least two types of ‘spirantisation’: (i) cases which involve |spread| in the input, behave in a very similar way to the cases of affrication and actually do produce fricatives, and (ii) the Spanish type, which involves |voice| stops and produces |voice| approximants.

The generalisation seems to be emerging that segments such as /f^h, θ^h, x^h/ can only be produced in lenition following a previous diachronic stage of affrication. As we see in the Liverpool case, the diachronic innovations can be telescoped synchronically to produce a transparent variable spirantisation, but such a process is dependent on a preceding diachronic innovation of affrication. I draw out the implications of this below. First we need to consider how the scenario proposed here can be modelled phonologically, given the assumptions made for phonology in chapters 1 and 3. The representations in (3.61) show the elemental structure of the relevant segments. The processes involved will naturally need to be able to map one of these onto the other:



This scenario clearly involves two processes which seem rather different from each other. In what follows, I refer to them as ‘affrication’ and ‘deaffrication’. Deaffrication is

⁸⁵ I assume that the fricatives produced, for example by the HGCS₁ and Liverpool English, are ‘strident’. It seems clear that [x] and [f], at least, involve a high degree of noise.

unproblematic. Given the assumptions of the model, we can formalise deaffrication simply as the loss of an already present [occlusion] element from the make up of the affricate (following the analysis of debuccalisation in section 3.1.3.1.8, and in line with the ideas developed by Harris). This is a phonologically minimal process and the innovation is clearly not unnatural.

Affrication is more problematic, however for it seems that a [frication] element has suddenly appeared in the make-up of the affricate in the process $k^h \Rightarrow kx^h$. This seems clearly unnatural in that it looks as if it would require an arbitrary rule filling in the [frication] element into the stop to give the affricate. Discussing an example of this type of affrication process (the HGCS₁), Davis & Iverson (1995) sum up the problem well, using partly binary features rather than elements: “why should aspiration - a laryngeal gesture implemented by the feature [spread glottis] - have resulted in affrication? Monosegmental affricates are now generally taken to be represented by a sequential specification for continuancy i.e. [-continuant]/[+continuant] ... but this configuration has no direct or apparent relation to the features of the larynx” (1995, 113). The representation that they offer for an affricate ([[-continuant]/[+continuant]]) can be straightforwardly translated into the elemental representation used here ([occlusion|/frication|]). Davies & Iverson’s solution to the problem is complex and involves the fission of aspiration from the segment, spreading of place features onto it and refusion of the two segments into an affricate. This analysis is quite lengthy and it is not clear that the processes which they propose are at all natural, or ever attested in languages. There is an alternative. In fact, there is an entirely natural solution.

The ‘problem’ of affrication is a problem for the notion of naturalness in process innovation only because the naturalness is typically sought in the wrong place. It was noted in section 1.2.2.2 that ‘naturalness’ refers to both articulation and to the acoustic properties of the speech signal. Ohala (eg, 1974, 1992) has long stressed the role that acoustic misperception, or we might say, reanalysis can play in phonological change and I propose that this idea can provide a straightforward account for affrication. We noted in section 3.2.2.2.6.1 that the presence of [spread] in a stop leads to a particularly noisy release phase. This is ‘aspiration’. As we saw in section 3.1.3.1, it is a mistake to conceive of aspiration as simply a delay in the onset of voice because this implies that it is manifested by the absence of any sound. In fact voiceless aspirated stops are accompanied by considerable audible noise in their release phase. Affricates, too, are accompanied by considerable audible noise in their release phase. The acoustic effects of both kinds of segment (an aspirated [spread] stop and a [spread] affricate) are very similar (while the articulations that produce them are very

different) and all that is required for the diachronic innovation of affrication is the reanalysis of the (voiceless, high-frequency) ‘aspiration’ noise as (voiceless, high-frequency) ‘frication’ noise. The reanalysing innovator thinks that speakers are producing an affricate and begins to produce one themselves. This has the effect of introducing [frication] in a ‘non-genetic’, abrupt but natural and quantally minimal, innovation. As soon as the acoustically-driven innovation has occurred in this way an affricate is in effect available to the phonology of a language and can be ‘captured’ by the phonology and even be related to an underlying stop by a process.⁸⁶ The process would not be ‘articulatorily natural’ but it would still be ‘minimalist’ (even if it boils down to the synchronically arbitrary addition of a [frication] element) as long as it remains transparently acquirable.

The final crucial assumption for this picture of affrication is to account for the cases where segments are affricated or spirantised but where there is often little audible aspiration in [spread] stops. The generalisations that exist in this regard were discussed in section 3.1.3.1.9, where it was shown that [spread] stops are deaspirated in certain phonological environments, for example, word-final. The potential problem that this causes for the account developed here is that there is little aspiration there which could be reanalysed as frication. But this is not a problem. I argued in detail in section 3.1.2 that we need to recognise the psychological reality and unity of the segment. Given this, the simple assumption is that speakers generalise from occurrences of the segment where aspiration clearly leads to reanalysis to those cases where this might be less clear. We thus have a natural means of process innovation for both affrication and deaffrication, even though they are entirely dissimilar processes. They are dissimilar because they reflect their different types of diachronic origin: phonetic reanalysis and phonological element loss respectively.

This account may seem theoretically undesirable because it fails to unify the two parts of the process $k^h \Rightarrow kx^h \Rightarrow x^h$ but when it is allied with the other generalisations that we have described, it does make a number of empirical predictions as to what is a possible lenition. If we assume the account here and the general unavoidable principles of uniformitarianism and minimal quanta, the predictions in (3.62) are made:⁸⁷

⁸⁶ As noted in section 1.2.2.1, the standard assumption is that this reanalysis would occur cross-generationally, but if an adult speaker’s phonology is malleable, as we might expect if it is not formulated by a phonological UG, then we do not necessarily need to commit ourselves to this assumption.

⁸⁷ I leave aside here the question of which element takes on the role of head in the output. It seems that there are generalisations to be made, for example, the input /t^h/ is related to an output which has [spread] as its head, as we have seen, and as will be further briefly discussed in chapter 4. This may well be the wider generalisation, but I do not pursue this here.

(3.62)

- i) the only ‘voiceless fricatives’ which can result from lenition are /f^h, θ^h, x^h/
- ii) spirantisation to segments such as these can only occur:
 - a) in stops which are characterised by |spread|
 - b) through an affricate stage

Thus:

- iii) $k^h \Leftrightarrow x^h$ can occur endogenously in the history of languages
(through the trajectory $k^h \Leftrightarrow kx^h \Leftrightarrow x^h$)

- iv) $k^o \Leftrightarrow x^o$ cannot occur: it is not a possible naturally-innovatable endogenous process
(there is no way to introduce a |frication| element because there is no aspiration in the release phase of /k^o/ which could provide the acoustic basis for reanalysis)

This ties in with both the generalisations made in section 3.2.2.1 and with Kirchner’s generalisation: ‘strident’ fricatives such as /f^h, s^h, θ^h, x^h/ can only be ‘created’ from affricates. We turn to the second part of his generalisation in the next section. The predictions made here are empirical and could be falsified by the discovery of a process which creates true voiceless neutral fricatives from neutral stops. For the moment, we can note that where such stops undergo ‘lenition processes’ (as in the history of the Romance languages, for example), these are typically ‘voicing’ processes, not spirantisation. Given the argumentation in this section, I extend the analysis developed here to the cases of spirantisation in Proto-Iranian, Kannada and Irish cases which were discussed in section 3.2.1.

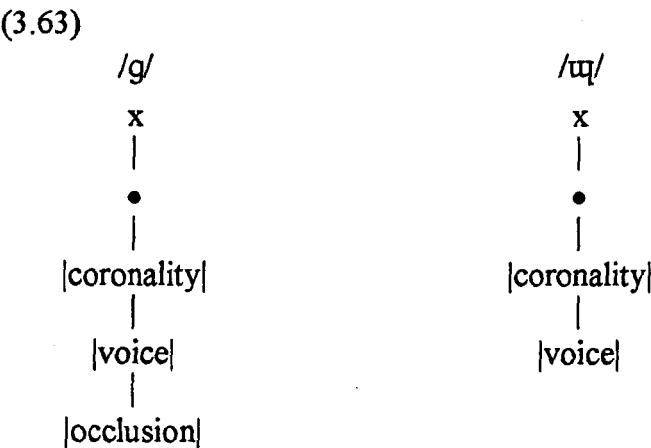
As we will note further below, this analysis also restores an age-old link between aspiration and affrication. The two have often been linked, with aspiration being seen as the ‘cause’ of affrication in some sense and even of spirantisation, but the mechanism through which this could occur has never been made explicit.

3.2.2.3.2 |voice| stop ‘spirantisation’

In the last section, we saw how many cases of spirantisation (those in spread languages) can be best accounted for, given the diachronic and synchronic phonological patterning. But there are other cases of lenition which have been described (elsewhere and earlier in this thesis) as cases of spirantisation. These include the case discussed from the history of Spanish in section 3.2.1.2. It was mentioned there, however, that recent detailed investigation (in Lavoie 2000, 2001) has shown that the Spanish lenition is not, in fact, a case of spirantisation at all. This is backed up by the second part of Kirchner’s (1998) generalisation, which was cited in full in the last section: “stops typically spirantize to weakly fricated or approximant continuants such

as [β, ð, γ] or [β, ð, ɰ]” (1998, 99). We have seen above that this need not apply to all cases of spirantisation, indeed, the Liverpool English segments are unmistakably fricatives.

The clear generalisation that is emerging is that only segments which include |spread| can truly spirantise; only they can be inputs to a process which has true fricatives as its output. Stops which are specified for |voice| cannot spirantise, but they can ‘approximantise’. Given the segmental representations which were adopted in section 3.1, this process is indeed minimal and phonologically very straightforward. It simply involves the loss of the |occlusion| element, as shown in (3.63). No |frication| element is introduced because the result is not a fricative.



It may well be that the |voice| specification is superfluous in the approximants. This would also allow the approximantisation analysis proposed here to be straightforwardly extended to ‘neutral’ stops which have no laryngeal element.

The predictions here are made explicit in (3.64):

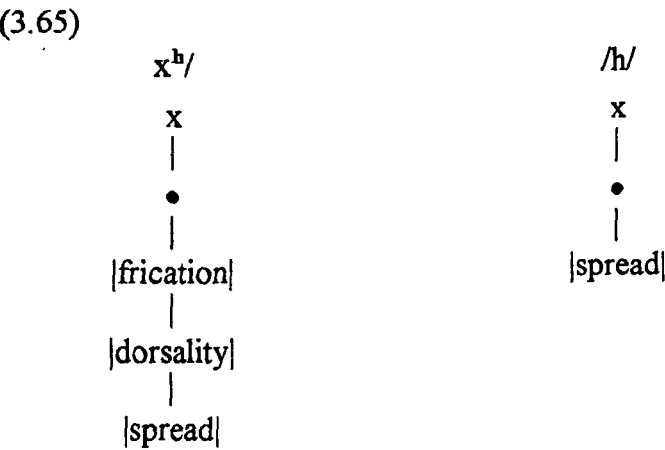
- (3.64)
- i) stops with |voice| (or no laryngeal specification) cannot spirantise to produce true fricatives
 - ii) the first quantum for such segments in lenition is to approximants

Thus:

- iii) g ⇒ ɰ can occur endogenously in the history of languages
- iv) g ⇒ γ cannot occur: it is not a possible naturally-innovatable endogenous process (there is no way to introduce a |frication| element)

3.2.2.3.3 Debuccalisation

We encountered cases of debuccalisation in section 2.1.1.5 (affecting the output of /x^h/, which had been derived in one of the GCS processes) and in section 3.2.1.4 (from Kannada /f^h/). We also briefly discussed another case from Spanish in section 3.1.3.1.8 and the analysis proposed there in terms of element loss still holds good. We saw there that the facts of fusion show that this debuccalisation leaves only |spread|. This means that the process involves the loss of two elements for a segment like /x^h/, as shown in (3.65):



This is essentially still the analysis of Lass (1976) and it may be that Lass is right to recognise (along with feature geometrists) some more structure in the segment than is shown here, differentiating between an oral ‘gesture’ or node and a laryngeal gesture. In terms of the possible predictions which are made by this analysis, it seems that only certain kinds of segments can debuccalise in this way. This is made explicit in (3.66).

- (3.66)
- i) only segments with |spread| in their make up can debuccalise to [h]

This seems right to the extent that debuccalisation to [h] is hardly ever reported for segments which are transcribed as voiced (for any case where it were to be reported, the analysis presented here predicts that it would be reanalysable in some other way).

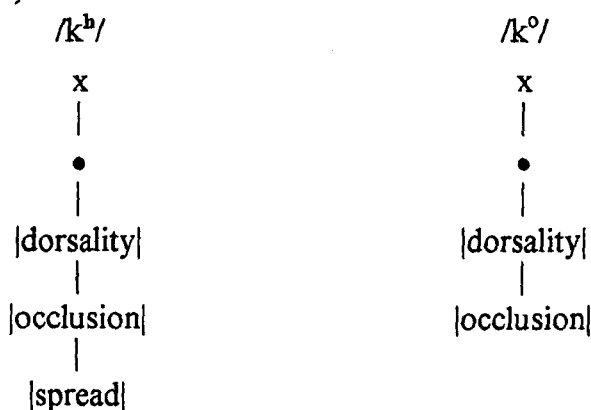
3.2.2.3.4 ‘Voicing’ and delaryngealisation

In the initial discussion of lenition in section 1.4, ‘voicing’ was portrayed as a simple case of ‘sonorising’ lenition. Some of the Germanic data in chapter 2 was described in these terms (the EIFV and the IGCW) and we also briefly investigated another case of ‘voicing’ from the

history of Spanish in section 3.2.1.2. In section 3.1.3.1, however, we saw that the traditional understanding of ‘voicing’ is over-simplistic. It is quite clear that not all processes which have been described as cases of voicing will necessarily be the same type of process. Because there are two phonological types of *T* and two phonological types of *M*, a process which has been described as producing an *M* from a *T* could really be a case $k^h \Rightarrow k^o$ or a case of $k^o \Rightarrow g$. Given that these are two different processes, it will not be surprising to see that they involve very different types of segmental effects.

The first of these, $k^h \Rightarrow k^o$ (or, for fricatives, $x^h \Rightarrow x^o$, of course) can in fact be recognised as a kind of opposite to the debuccalisation discussed in section 3.2.2.3.4. Whereas debuccalisation involves the loss of oral articulation, the change under consideration here involves the loss of laryngeal articulation, as is shown in (3.67).

(3.67)



The potential for this type of process has not previously been recognised, but it will be clear that the possibility of such processes is predicted by the phonological model adopted here. I call it ‘delaryngealisation’, on the model of the term ‘debuccalisation’ (which is also known as ‘deoralisation’). Delaryngealisation is, in fact, the obvious companion that we might expect to contrast with debuccalisation because the former involves loss of constriction in the larynx but retention of constriction in the oral cavity and the latter involves loss of constriction in the oral cavity but retention of constriction in the larynx. This type of analysis also extends to the other types of segments discussed in terms of their laryngeal specifications in section 3.1.3.1. For example, the change from ejective to neutral stop is a form of delaryngealisation (it involves the loss of |constricted|). The particular case of delaryngealisation shown in (3.67) has also been referred to as ‘deaspiration’.

This will not account for the Spanish case of ‘voicing’ however. Spanish is, and probably always has been, a |voice| language. This means that any ‘voicing’ will involve the process

$k^{\circ} \Rightarrow g$, because the *Ms* in Spanish are characterised as $|voice|$ segments and this will involve the previously neutral $/k^{\circ}/$ picking up $|voice|$. It is notable that in the Spanish case from section 3.2.1.2, this $|voice|$ specification is clearly spread from neighbouring voiced segments (in traditional terminology, it is intervocalic voicing), and it seems that this is typically the case for this type of ‘sonorisation’. I discuss the implications of this below.

The possible predictions from the model in terms of the processes discussed in this section are not so clear, but we might nonetheless expect $k^h > g$ to be impossible in one quantum, and we might hypothesise that $k^{\circ} > g$ needs a local source of some sort in order to be able to acquire $|voice|$.

3.2.2.3.5 Summary and prospect

In this section, we have considered in detail each of the types of obstruent lenition which we started with in section 1.4. In doing so, we have seen that there are, in fact, more types of process than was recognised there, and this refinement in the understanding of what types of phonological processes are actually cases of ‘lenition processes’ can be seen as a distinct increase in clarity. Not only have we recognised that there are more types of process than were previously recognised, but we have also seen that several of the different process-types pattern differently phonologically.

As we saw briefly in section 1.4 and in quite some detail in section 3.2.2.2, much of the lenition literature has attempted to recognise a unity among lenition processes, so that there is some way of identifying what counts as a case of lenition, and what is not. We saw that there have been several attempts to formally unify the processes in phonological models, but that none of these quite succeeded. In the current section (3.2.2.3) we have seen that several different types of process seem to be involved in ‘lenition processes’. One of the cases touched upon above does not seem so clearly to belong with the others. While most of the processes described in this section do not seem easily characterisable as cases of well-recognised types of phonological change, such as epenthesis, dissimilation or metathesis, one of them is. The Spanish-type of intervocalic voicing seems to be a case of assimilation. We saw in section 3.2.2.2.2 that some authors have tried to class all cases of lenition as simply a sub-class of assimilation, but that attempt at classification was shown to fail. If we are to seek a principled classification, however, it will be wise to exclude those processes which already belong to one classification, so we can detach the voicing to $|voice|$ from the other cases discussed here, for the others are clearly not assimilations. This type of process already

belongs to a taxonomic class, so it would be wrong to seek to include it with 'lenition', if lenition is to be a different class of process.

We seem to be left with a situation where lenition cannot really be thought to exist at all. We have seen that there is no formal unity in the processes involved. While approximantisation, debuccalisation and delaryngealisation can be seen as simple cases of element loss (as in Harris's account from section 3.2.2.2.6.1), affrication and spirantisation are inextricably linked with acoustic reanalysis. If there is no coherent definition which can link these processes then, strictly speaking, there is no such thing as obstruent lenition.

There is one way, however, that we can group the processes which allows for a certain insight into what it is that they have in common. This is certainly not a highly 'phonological' definition because it does not derive from any primitive of a phonological theory. Rather, it is a 'historical phonological' definition because it relies on factors that are relevant to the innovation of 'lenition processes'. It relates to synchrony to the degree that diachrony is normally related to synchrony, which is through the constraints that a consideration of process innovation can place on the nature of synchronic processes. In section 1.4, we saw that there have been attempts to define lenition in terms of a specific set of environments in which it occurs. As we have seen, this approach fails because they are very different.

There certainly are generalisations that can be made about the connection between lenition and the environments in which it occurs, and we turn to these in the next section, but as we have seen, there is not a simple set of 'lenition environments' which would allow us to define lenition in terms of them. Obstruent lenition seems to be comprised of a set of processes which can be innovated in historical phonology to 'affect' these obstruents without changing their place of articulation (that is, without changing their place element). In section 1.2.3, we saw that a distinction can be made among three key types of process in terms of the change that they effect in a segment. These are (i) 'conditioned changes', which can be thought in some way to be directly 'caused' by the environment in which they occur due to the influence of adjacent or closely neighbouring segments, (ii) 'weakly unconditioned changes', which are not entirely context-free in that they may not affect every occurrence of a segment in a language, but which cannot be clearly seen to be 'caused' by the environment in which it occurs, and (iii) 'strongly unconditioned changes', where phonological environment played no role at all.

The best understanding of the type of processes discussed in this thesis comes in the simple realisation that lenition is a cover term that we can use to group together those processes which effect weakly unconditioned change in obstruents which do not alter their major place of articulation. We will see in the next section that lenitions can be affected by their prosodic

and melodic environment, but, I argue, they are not caused by it. Lenition clearly tends to affect classes of segments, but as the Kannada example and the ‘GCS’ debuccalisation show, this is not necessarily the case. The kind of definition adopted here means that we should not expect there to be an underlying unity in terms of the phonological characterisation of the processes that are involved.

When discussing the notion that certain phonological environments can be seen to ‘cause’ phonological processes (in conditioned changes) whereas others do not (unconditioned changes), it is important to bear in mind the fact that the notions of causation and explanation are complex ones in historical phonology (as they are elsewhere, too). This applies equally to the idea that certain phonetic and phonological properties which are inherent in the segments can ‘cause’ processes in unconditioned changes. These points were discussed in section 1.3.1, and, in line with the points made there, it will be clear that if we say, for example that a |spread| specification can ‘cause’ affrication, then the notion of ‘cause’ invoked here is only partial. Not all |spread| stops affricate all the time, so we can see that, while |spread| is a necessary part of the causation, it is not sufficient by itself. What is needed as well, is the chance effect, whereby speakers act (unconsciously) on the potential which is offered by the linguistic system. Naturally, this is the same for the innovation of any process in historical phonology, and is not restricted to lenition. It is the general failure of the notion of deductive-nomological explanation which we noted in section 1.3.1, following Lass (1980). As we saw there, this illustrates the fact that, while we can make certain falsifiable predictions in connection with the account proposed here, we cannot use it to predict absolutely which ‘route’ a language will take. Another possibility for a |spread| stop is to lose its laryngeal element, as we saw in section 3.2.2.3.4.

As we saw in section 1.3.1, while we should not give up the goal of formalising falsifiable predictions, it is not appropriate to expect these to be absolute predictions which could result from a statement of absolute, necessary and sufficient causes. These caveats apply to the way in which the term ‘cause’ is used in this section and elsewhere (along with semi-synonyms such as ‘account for’, ‘be responsible for’, and others).

The understanding of lenition which is developed here means that the notion of segmental ‘strength’ which has been developed in connection with lenition trajectories (as discussed in section 1.4 and elsewhere in this thesis) can only be understood as a metaphor. This does not make it a useless concept. Metaphors can and do play an important role in the development of academic disciplines and should often be recognised as helpful, insightful notions which can help us to the ‘relief from puzzlement about some phenomenon’ which is a somewhat explanatory goal, as we saw in section 1.3.1. Both Fónagy (1963) and Lass (1997) (and see

their references for copious other work on the topic) show how metaphors can play and have played an important role in shaping of thinking in linguistics, to an extent that it can hardly be denied that they have increased our understanding.⁸⁸ In a slightly different connection, Wells (1987) writes, on the place of metaphors in linguistics, that “we should be neither misled by metaphors nor frightened of them” (1987, 42; cited in McMahon 1994), and it seems reasonable that the concept can be used to form ‘strength scales’ and ‘lenition trajectories’.⁸⁹

There is certainly no reason why we should not use the term ‘lenition’ to group the kind of processes that have been covered in this thesis, as long as it is understood in the way explored in this section. For the construction and understanding of lenition trajectories, we should recall the starting point for these in section 1.4. We saw there that a common beginning for the discussion of segmental strength and weakness is normally Vennemann’s personal communication to Hyman (1975, 165) that “a segment X is said to be weaker than a segment Y if Y goes through an X stage on its way to zero.” This definition, in fact, seems exactly right. It relates the notion of segmental strength to the diachronic dimension, where it belongs. It can allow us to create ‘lenition trajectories’ of the type that have been referred to at various places in this thesis. We can take Vennemann’s observation as the only reasonable basis for this, along with Lass & Anderson’s (1975) initial observation that lenitions involve “sequences of changes that tend to repeat themselves again and again in the histories of languages” (1975, 150).

Given the understanding of phonology and of the patterns that can be seen in the innovation of lenition processes, we can construct a lenition trajectory such as that in (3.68), which is again exemplified for velars and where ‘ \Rightarrow ’ indicates a ‘lenition process’, apart from the case in the box, which, as we saw above, is an assimilation, and is included to allow all the processes discussed to be treated together - it forms a ‘bridge’ on the lenition trajectory.⁹⁰

⁸⁸ Fónagy (1963) shows in some detail that metaphors have always been present in phonetic and phonological thinking, to a greater or lesser extent.

⁸⁹ Lass (1997, 41) writes that “metaphorical objects (if we want to call them that) live in a rather special half-world of their own; we may not be able to pin down what they are, but they’re so useful that it would be counterproductive to get rid of them. They are at least something more than decoration.” This seems to me to be entirely applicable to the notion of ‘segmental strength’.

⁹⁰ This ignores the status of some of the segment types which were discussed in section 3.1.3.1, such as ejectives and implosives, which as we have seen, can debuccalise, and can also undergo other lenition processes. See Fallon (1998) for discussion of the lenition patterning of such segments.

(3.68)

$$\begin{array}{c} k^h \Rightarrow kx^h \Rightarrow x^h \Rightarrow h \Rightarrow \emptyset \\ \Downarrow \\ \boxed{k^o \Rightarrow g} \Rightarrow \text{ɥ} \end{array}$$

The understanding of lenition which has been developed in this section seems to be the only coherent one. We have seen that among ‘lenition processes’, there are different types of processes and so the underlying unity among them is to be sought in the patterning in their diachronic innovation. This may not be entirely theoretically satisfying, but it seems the right conclusion. There are numerous statements in the literature along the lines of Crowley (1992) “[t]he concept of lenition is actually not very well defined, and linguists who use the term seem to rely more on intuition or guesswork than on detailed understanding of what lenition is” (1992, 39). It seems to me that this confusion is due to a failure to recognise the kind of similarity that exists among the processes. A range of factors are responsible for ‘lenitions’ and they involve different kinds of phonological process-types. What they have in common is that they are weakly unconditioned, that is, they are not ‘caused’ by their environment, but can be constrained by it, as we will see in the next section. In a slightly different context, but nonetheless relevantly, McMahon (2000b) has described a problem that is sometimes encountered by accounts for linguistic phenomena that are not theoretically elegant: “relinquishing a single explanation type ... can be seen as admitting defeat; it might also acknowledge a lack of ingenuity, since ... ‘plausible stories can always be told’... however, concocting plausible stories might not always constitute progress” (2000b, 182). It seems to me that, while many have tried to construct plausible unitary stories for lenition, we need to recognise both the diversity of process types, as well as a kind of unity, if we are to explain or understand the concept.

3.2.3 Lenition and phonological environments

The definition provided for lenition in the last section focuses on the types of process involved and on their diachronic innovation. Bearing in mind the caveats on the word ‘cause’ which are clear from previous discussion of explanation and causality in this thesis, we can note that lenition processes, unlike certain other process-types, are not ‘caused’ by the phonological environment in which the segments occur. Nonetheless, because they are not strongly unconditioned processes, they still have the potential to be affected by their environment.

In this section, I discuss the interaction that can occur between lenition processes and phonological environments. We will see that the standard way of conceiving of the connection is not the most insightful. I show that most previous work on the issue of environments has focused only on prosodic factors and has missed the recognition of generalisations that can be made concerning the effect of melodic factors.⁹¹ Perhaps more profoundly, I propose that the interaction between lenition and environment is not best thought of in terms of the idea that certain environments promote the innovation of lenition processes, but rather that certain environments can inhibit it. Given the focus on melodic factors here, I will have little to say on the interaction between prosody and lenition. In general, this seems much better understood, and I do not seek to make a contribution to debate in this area. Some of the melodic effects to be discussed here have been described as cases of ‘geminate inalterability’. I show, however, that the way in which such factors have typically been interpreted is at fault and that previous discussion has missed certain generalisations.

Most of the lenition processes that we have discussed so far, especially the cases of affrication, spirantisation, approximantisation and debuccalisation have shown environmental effects, in line with their status as only weakly unconditioned processes. In order that we might come to understand these effects, it will be worth considering the precise nature of the effects for each of the cases of lenition that we have considered here. I do this in section 3.2.3.1 below, where I also seek to draw out any clear generalisations. At the start of this section, I briefly discuss what is the best way to approach the data and the generalisations. In section 3.2.3.2, I present a novel means of understanding the melodic generalisations that are observed in section 3.2.3.1 and tie these in with existing ideas about the effects of prosodic environments.

3.2.3.1 Environmental observations and generalisations

In section 3.2.1.1, we considered a set of phonological environments, many of which had previously been claimed to be relevant to the interaction between lenition and environment in previous discussions in the literature (for example, Escure 1977, Harris 1994, 1997, Ségéral & Scheer 1999). These are reproduced here as (3.69):

⁹¹ There is less discussion of previous proposals from the lenition literature in this section than there was in the last. This is despite the fact that some of the approaches discussed there are connected with ideas which are intended to account for environmental interactions. If a proposal was shown to fail in the last section on the grounds that it does not achieve what it set out to do in terms of the segmental aspect to lenition, then its environmental aspects need hardly be discussed in this section. Certain aspects of such proposals are discussed here, however, where they are particularly relevant.

(3.69)

- | | |
|--|----------------------------------|
| A [__#] - 'word-final' | |
| B [__c] - 'coda' or 'pre-consonantal' | # = word boundary |
| C [v__v] - 'intervocalic', 'medial' | c = any consonant |
| C1 [\acute{v} __ (v)] - 'foot-internal', 'post-stress' | v = any vowel |
| C2 [(v) __ \acute{v}] - 'foot-initial', 'pre-stress' | \acute{v} = any stressed vowel |
| D [c__] - 'onset' or 'post-consonantal' | |
| E [#__] - 'word-initial' | |

Several of these environments did indeed prove to be useful in the description of the environmental patterning of the lenition processes that have been discussed in this thesis. Some of these are clearly prosodic environments, and we saw that environment D [c__], while it can have a prosodic interpretation, also shows a range of effects according to the melody of the 'c' involved.

This has not consistently been recognised previously and has not received the consideration that it warrants. For example, Escure (1977), basing her generalisations on a consideration of certain types of lenition, writes that lenition "is most likely to affect clusters ... and single consonants in utterance-final position and least likely to occur in utterance-initial position" (1977, 58). This misses crucial generalisations as to the nature of the clusters involved, however. Many authors (eg, Hayes 1986, Schein & Steriade 1986, Elmedlaoui 1993, Kirchner 1998, 2000) have recognised that geminates are particularly resistant to lenition processes, and these are clusters of a sort, in that they involve two timing slots in separate syllables, as shown in the discussion of prosody in section 3.1.4, but there is a little more to the interaction between melody and inhibition than this.

The approaches to geminate inalterability have also been described as 'blocking' effects (for example in Kirchner 1998, 2000 and Bermúdez-Otero 2000) with the implication that a process was introduced with general applicability, but was inhibited from applying in certain environments (ie, in geminates).⁹² This contrasts with an approach to the innovation of lenition processes which is perceivable in other work, which claims that some of the environments in (3.69) actually serve to promote lenition (as explained in section 3.2.1.1, these are generally taken to be A, B and C1).

There seems thus to be a conflict between approaches which recognise 'lenition

⁹² Another terminological tradition describes similar effects as 'protection' (for example, Lass & Anderson 1975).

promotion' as a possible factor affecting the innovation of lenitions and those which recognise 'lenition inhibition'. In fact, both of these could be extended to cover all environments for any lenition. We could see those environments where lenition typically occurs as 'promoting' environments and those where lenition does not typically occur as 'inhibitory' environments. These are really two sides to the same coin, though, and we only need to describe one of the two in order to fully describe the environment of a lenition, because the other follows as its obvious opposite. It is unparsimonious to have two sets of generalisations which account for the same thing, so it seems fair to reject one of these two ways of viewing the environmental interaction and to work with the other.

Using slightly different terminology, Ségéral & Scheer (1999) propose that the environmental phonological patterning of lenition should be described purely in terms of inhibition. They point out that the fact that languages change is one of the few absolute universal characteristics of language. In terms of phonology, one of the things that this involves is the introduction of new processes, as was discussed in chapter 1. The introduction of new phonological processes is thus not a surprising fact. Rather, it is almost to be expected. It is indeed perfectly natural to innovate new phonological processes, and it arguably becomes surprising, once a process is introduced, that it does not occur across the board. Given the potential diachronic effect of the type of phonetic and phonological factors which have been discussed in this thesis (in chapter 1 and in section 3.2.2.3), we do not need to account for the innovation of processes which internalise them by using the notion of 'promoting' environments because we already have an account of the promoting factors. The interesting environments are likely to be those which inhibit phonological processes. I thus focus on these, partly following Ségéral & Scheer (1999).

This approach inverts the conventional view (of 'promotion') and opens up a different and prospectively more fruitful perspective - the consideration of what prosodic and melodic factors prevent the onset of a process. Once a full description of these environments is given, then the notion of a 'promoting environment' simply becomes 'those environments which are not inhibitory'. For our purposes, it thus becomes important to consider where lenitions do *not* typically occur, that is, traditionally formulated, the exceptions to processes.⁹³ The focus on 'exceptions' is not new, of course, but has not always received the consideration that it deserves. I turn to the ways in which inhibition can be best formalised in the next section. In the remainder of the current section, I return to the lenitions that we have encountered in this

⁹³ In an early, and influential piece of work which focused on of exceptions, Verner (1876) wrote "there must be a rule for exceptions to a rule".

thesis and consider them one by one in terms of the patterns of inhibition that they show.⁹⁴ I later consider what generalisations arise, both in terms of prosody and melody.

The processes in the GCS included two clear lenitions: GCS₁ and GCS₂. The philological description of these in section 2.1.1 showed that the majority opinion for GCS₂ is that there was no environmental effect and that the process occurred in all environments. I accept this here, but return to the issues in section 4.1.1. There are clear 'exceptions' to the GCS₁, however, and, given the discussion of such issues in section 1.2.2.1, we can take these as a fossilisation of the environment of the original lenition process. We saw in the discussion in section 2.1.1 that the alveolar input did not lenite in clusters following a stop, and that none of the inputs lenited following the alveolar fricative. In section 3.2.1, we also encountered this inhibiting environment, and it was formalised as [s__]; I continue this practice here. There were no prosodic effects for the GCS₁.

The inhibition shown in the HGCS is similar to that of the GCS but somewhat more complex. The lenition process here is the HGCS₁. As we saw in section 2.1.2 and in the discussion of lenition processes in section 3.2.2.3, the lenitions involved affrication in certain environments and, in a subset of these, spirantisation. Section 2.1.2 included substantial philological detail about the patterns of inhibition of the HGCS₁, showing dialectal differentiation. In all dialects, the lenitions were inhibited in the environment [s__]; also, the lenition of /t^h/ was inhibited in clusters preceding an /r/ and following a fricative. In environment E [#__] and other cases of D [c__] (ie, apart from those just discussed) the situation varied according to dialect. We can also note that the least inhibitory environments seem to be A [__#] and C [v__v], where spirantisation has occurred.⁹⁵

The full details are given on the chart from Keller (1978) in (2.13). Here, I illustrate some examples. In many varieties, E [#__] was entirely inhibitory for certain segments, at least, so for example, in Mosel Franconian, Rhine Franconian, Ripuarian and Low Franconian, neither /p^h/ nor /k^h/ affricated at all in E and in Low Alemannic, /k^h/ did not affricate in E. In other varieties (Highest and High Alemannic), E [#__] did not have this inhibiting effect. The case of D [c__] is even more intricate. It can be seen to have been generally inhibitory, for example in Ripuarian, lenition of /k^h/ and /p^h/ was inhibited here generally (where the consonant was /r, l/ or a nasal or the 'first half' of a geminate). The same is true for Low Alemannic but only for /k^h/). In Mosel Franconian and Rhine Franconian, however,

⁹⁴ Not all of the examples discussed here show clear evidence for any inhibition, for example the EIFV, IGCW and the Greek cases from section 3.2.1.6.

⁹⁵ For A [__#], this was not true when the segment was preceded by a consonant, but these cases can better be seen as examples of D [c__], the effects for which are explained below.

environment D [c__] was only inhibitory (for /k^h/ and /p^h/) where the consonant was a nasal or in a geminate. In High Alemannic, lenition of /k^h/ was only inhibited when preceded by a nasal. The generalisations that seem to emerge from the fine patterning of the processes in the dialects of German are that E [#__] can certainly be inhibitory and that, while D [c__] can be inhibitory as a unified environment, this is sometimes only the case where the consonant is a nasal (in section 3.2.1, this was described as [N__]) or in gemination.

The Spanish approximantisation data from section 3.2.1.2 shows certain definite similarities to the patterns of inhibition shown by the HGCS₁. Approximantisation occurs straightforwardly in environments A [__#] and C [v__v] (like HGCS₁ spirantisation), as well as in B [__c]. Environment E [#__] was shown to be inhibitory by itself, but not when non utterance-initial and preceded by a vowel (ie, [v#__]). Environment D [c__] was found to show an interesting pattern of inhibition, as in the HGCS₁ and the GCS. In Spanish, the D-type environment [N__] is inhibitory and also, /d/ does not lenite after /l/, whereas /b/ and /g/ do approximantise in this environment.

The situation in Liverpool English is again rather similar, as can be seen from the data in section 3.2.1.8. In environments A [__#], B [__c] and C1 [v'__ (v)] lenition to fricatives is common. In environments E [#__] and C2 [(v)__ v'] lenition to affricates is common. In environment D [c__] the inhibitory effects of melody again become apparent. For both /t^h/ and /k^h/, the environment [s__] is completely inhibitory and in [N__] affrication can occur, but spirantisation cannot. For /t^h/, the environment [l__] only allows affrication, whereas [l__] allows affrication or spirantisation for /k^h/ . It seems further that [f^h__] and [k^h__] inhibit lenition for /t^h/.

In the Irish historical lenitions, which were described in section 3.2.1.5, following Thurneysen (1909), lenition was inhibited in environment E [#__] when preceded by a vowel, in a similar way to the Spanish case (ie, [c#__]). Also, the environment D [c__] show some notable melodic effects. The processes were inhibited in geminates and in [s__] and in [N__]. There were also certain other inhibitory effects, as Thurneysen explains: for stops after /r/ and /l/, and the alveolar /t^h/ did not lenite after /x/, and, finally, /b, g/ did not lenite after /ð/ (which developed from /z/).

There are three other cases among the lenitions from section 3.2.1 which illustrate inhibition. These all show melodic effects. Proto-Iranian was seen in section 3.2.1.3 to show

the inhibitory effects of [s__] and [N__], and both the Kannada case from section 3.2.1.4, and the Bantu case in section 3.2.1.7 show the inhibitory effect of [N__].

There seem to be several clear generalisations that can be made about the patterns of inhibition shown in the data discussed here. These can be divided into a set of prosodic generalisations and a set of melodic generalisations, which typically hold for the environment D [c__]. The prosodic generalisations are that E [#__] and C2 [(v)__ ´] can be inhibitory, and it is possible that the prosodic interpretation of D [c__] (when it stands for an onset) can be inhibitory, too. Environment E was inhibitory in Liverpool English, and the HGCS₁ in that spirantisation was inhibited, although affrication is not. When utterance initial, E was shown to be inhibitory in Spanish and Irish. Environment C2 seems to be inhibitory in Liverpool English, and it is possible that the onset nature of D played a role in some of the dialectal generalisations discussed above for the HGCS₁.

The clear melodic generalisations for D [c__] are that full geminates can be inhibitory, as can [s__] and [N__]. Geminates were at least partially inhibitory in the HGCS₁, as was the environment [N__], as shown in some detail above. The environment [N__] was also inhibitory in Spanish, Kannada, Bantu, Irish, Proto-Iranian and Liverpool English. The environment [s__] was inhibitory in the GCS, the HGCS, Proto-Iranian, Irish, and Liverpool English. There also seem to be other possible effects, as in Irish, for example, but there do not seem to be clear generalisations behind these.

As has been made explicit, the approach adopted here assumes that the environmental effects that can be observed in lenitions are most insightfully characterised by assuming that the phonetic and phonological factors which were described in section 3.2.2.3 are responsible for the innovation of the processes involved. As we saw in that section, lenitions are ‘unconditioned’ processes because they are not ‘caused’ by their environment. On this account, once they have been innovated, they are free to affect all occurrences of a segment unless this is inhibited. We have seen in this section that certain generalisations can be made about what kind of environments are inhibitory. What remains is to consider why they might be so. I turn to this in the following sections.

3.2.3.2 Prosodic lenition inhibition

The generalisations that were recognised in section 3.2.3.1 for the prosodic inhibition of lenition are widely shared in the literature. It seems that ‘initial’ prosodic environments are inhibitory: there is evidence, backed up by other work, (for example, Escure 1975, 1977, Harris 1990, 1994, 1997) that a segment in word-initial or foot-initial environment is often ‘protected’ from the effects of a lenition process. There have been several theoretical proposals which seek to account for these effects. As indicated in the introduction to this section, my main aim here is to investigate and account for the less well recognised effects of melodic inhibition, so I do not investigate the prosodic proposals in detail here, but I will make use of the generalisations of prosodic inhibition below, so some words are in order.

Escure (1977) presents a hierarchy of environments which are more or less likely to allow (or inhibit) lenition. As we saw in the short quotation from her work above, this ranges from utterance-initial, through word initial, intervocalic and coda-like environments to word and utterance final. The presentation and consideration of the hierarchy is essentially the same as that which we discussed for scales of individual segmental strength in section 3.2.2.2.1, and while it describes key generalisations (although the whole of the hierarchy is not backed up with evidence), it does not seem to provide any kind of explanatory insight, and it differs little from simply stating that word-initial and foot-initial environments are inhibitory.

Much of Kirchner’s (1988, 2000) approach to the interaction between environment and lenition focuses on geminates, and I deal with this in the next section, but, as we saw in section 3.2.2.2.3.1, Kirchner also uses a (potentially very large) number of ‘positional faithfulness’ constraints to account for lenition inhibition (which he calls ‘blocking’, as we saw above). These can be ranked above LAZY if a lenition shows inhibitory effects. For example, Kirchner (2000) uses IDENT (cont/onset) and IDENT (F/#__). Respectively, these can be understood as ‘a segment’s feature value for continuant may not be changed from the underlying to the surface form when the segment involved is in an onset’⁹⁶ and ‘a segment’s value for feature ‘F’ (which could be [continuant] or [voice], for example) may not be changed from the underlying to the surface form when the segment involved is in word-initial position’. The formulation of a set of constraints such as these may allow a precise description

⁹⁶ This kind of phrasology would not be used in OT, of course. A more OT-true gloss would be ‘a segment’s feature value for continuant must be the same in the input and the output when the segment involved is in an onset’. The basic meaning of the two glosses is the same, however. In fact, without the more clearly derivational terminology that I have adopted, and with the OT assumption of ‘richness of the base’, Kirchner is forced to assume a set of ‘fortition-inducing’ constraints, as was also briefly mentioned in section 3.2.2.2.3.1. One of these is (*[+cont, -son]/#__). This translates as ‘in obstruents, whether the input is a fricative or a stop underlyingly, make sure that it is a stop in word-initial position’. While the need for such constraints is indeed a consequence of the assumption of ‘richness of the base’, they do not seem a particularly insightful approach to the situation.

of the environments where lenition is inhibited, but, rather like Escure's hierarchy of environments, they seem to be a simple restatement of the description, rather like simply noting that, for example, word-initial and foot-initial environments are inhibitory.

Harris (1994, 1997) has a rather different approach.⁹⁷ This is connected with his description of lenition as element loss which was explored in section 3.2.2.6.1. Harris ties in the environmental effects which can be observed in lenition with the notion of phonological licensing, which he traces back to McCarthy (1979) and Goldsmith (1989), among others. The fundamental idea is that all segments, and also elements below the level of the segment, must be licensed if they are to exist. Certain segments are directly licensed (for example, the stressed vowel which is the head of a foot). Harris develops a notion of 'licensing inheritance' through which all the other segments in a phonological domain receive their necessary licensing. This involves licensing potential being passed on from skeletal slot to skeletal slot, decreasing in power as it goes. Those positions which receive decreased licensing power are those positions where lenition is typically not inhibited. This idea may provide a key to understanding prosodic inhibition, to the extent that we might think that only certain prosodic positions receive the necessary licensing to inhibit the innovation of a lenition process.

Whatever the precise mechanism, it seems clear that 'initial' environments can be prosodically inhibitory to lenition. Because my aim is to focus on melodic inhibition, I do not formalise the process through which this prosodic inhibition occurs, but simply note it, in a similar way, in fact, to certain other approaches to the issue, as shown above. It could well be that prosodic licensing is responsible for the inhibition, or it may simply be a facet of the greater prosodic prominence which initial segments have over medial or final prosodic positions. I turn now to the factors relevant to melodic inhibition, as, it could be argued, these may well reveal interesting facts concerning the nature of the phonological interaction between segments.

3.2.3.3 Melodic lenition inhibition: melodic molecules

The generalisations observed above concerning melody and inhibition were that geminates can be inhibitory, as can certain other D [c__] environments, including [s__] and [N__]. It seems that in certain cases, the melodic material which is connected to the preceding skeletal slot in a word (in line with the type of representations which were exemplified in section 3.1.4). One crucial aspect of the observation concerning the environment [N__] needs to be

⁹⁷ Some other approaches to the issue which rely on somewhat different theoretical constructs, but are not entirely dissimilar in principle are found in Ségéral & Scheer (1999) and Dienes & Szigetvári (1999).

made explicit here. The nasal referred to here in this environment does not represent just any nasal. In all the examples given above, the second, crucial part of the observation is that the nasal is homorganic with the following segment (in which lenition is inhibited). Such clusters are sometimes referred to as ‘partial geminates’ (eg in Harris 1997, Kirchner 2000).

The generalisation concerning the inhibitory effect of geminacy has been known as ‘geminate inalterability’ since Hayes (1986). In a clear overview, Elmedlaoui (1993) shows that previous work on the issue goes back to Guerssel (1978), at least. Much of this work has been conducted in frameworks which use nonlinear phonological representations of the type adopted here in section 3.1.4 and elsewhere. These allow a representation of geminates along the lines of that which was given in sections 3.1.5 and 3.1.4.2. It is repeated here as (3.70) for convenience:

(3.70)



Hayes (1986) accounts for the fact that geminates do not tend to undergo lenition processes by proposing a ‘Linking Constraint’ which relies on the existence of the association lines which link elements (which are features for Hayes) to root notes, and hence to skeletal slots. The ‘Linking Constraint’ states that “[a]ssociation lines in structural descriptions are interpreted as exhaustive” (1986, 113). This is formulated as a constraint on rules and has the effect that rules (such as spirantisation rules) which are formulated to affect single segments do not also affect any part of a geminate, even though the structural description of a rule might include the first half of a geminate; because of the representation of geminacy as the sharing of elements (or features), with multiple association lines, the rule fails to apply. The original insights behind the idea are valid, and have been extended to include the inhibition in partial geminates. Other similar accounts have been proposed which seek to interpret the inalterability effects in various ways, but such accounts typically have a clear problem. As Elmedlaoui explains, they “all share tacitly ... the assumption that a given spirantization rule restricted so as to apply only to geminates should not be less natural than one which is restricted so as to apply to just simplex segments. These two kinds of rule are equally

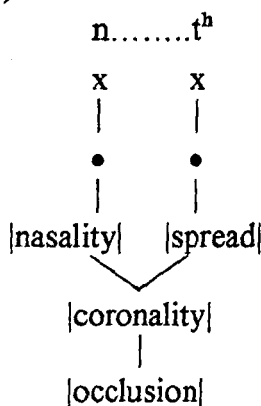
conceivable in view of those proposals, and are tacitly assumed to be equally operative and likely to take place” (1993, 134). Given the way that such constraints as Hayes’ ‘Linking Constraint’ are formulated, the inhibiting effect derives from a condition on the process, not from any property of the geminates or partial geminates themselves. As Elmedlaoui (1993) explains, given this state of affairs, the model predicts that there will be some lenition processes which only affect geminates, just as there are other processes which only affect non-geminates because both types of rule are just as easily formalisable given the assumptions of the theory. This is not really an empirical hypothesis, as we saw in section 1.3.1, because it is a positive ‘strong’ prediction, but nonetheless, it is quite problematic for this position that, as Churma (1988) notes, for all the lenition processes that we know of, “[n]o spiranitzation rule is restricted so as to apply only to geminates” (1988, 3, cited in Elmedlaoui 1993). The fact that the unavoidable prediction from the model is unfalsifiable and that it has received absolutely no back-up in the form of such processes suggests that we should search for an explanation elsewhere.

I propose that this problem arises in previous analyses because the cause of melodic inhibition has been sought in the wrong place. We can recognise a better generalisation which will also allow us to widen the generalisability of the claim; I explain this generalisation in the remainder of this section.⁹⁸

It is evident from the discussion here that some degree of element sharing is involved in the environments which show these melodic effects. For the full geminate in (3.70), this is clear, but it is also true for partial geminates, as shown in (3.71). This is a representation for the cluster /nt^h/; it shows that the cluster shares both place and |occlusion| elements, like all such [N__] clusters.

⁹⁸ Kirchner (1998, 2000) seeks to find an explanation in terms of the ease of articulation, because geminate fricatives would cost more effort than stop geminates, but this lacks the ability to generalise in the way developed below.

(3.71)



It seems that the sharing of elements is crucial. This simple generalisation also accounts for an asymmetry that is found in the Spanish data from section 3.2.1.2. We saw there that a homorganic preceding nasal inhibits lenition, and also a preceding segment /l/ is inhibitory for /d/ but not for /b/ or /g/. This fits with the ‘sharing’ generalisation because precisely /ld/ share a place element, whereas /lb/ and /lg/ do not. The recognition by Hayes (1986) and others of this generalisation is important, but their formalisation has been shown to be at fault. We need another way of conceiving of the melodic effects.

I propose that this sharing of subsegmental material gives a segment ‘strength’ by locking it into the word’s phonological structure. It is this strength-through-sharing which inhibits the lenition process from affecting the segment. The basic idea is that the two segments in a partial or full geminate⁹⁹ are bound together through element sharing and hence fixed into the phonological structure of the word. This fixing is the source of the strength that is given to the segment which enables it to resist the onset of the lenition process.

The notion of ‘strength’ developed here is very different to the type of segmental strength which was discussed in sections 1.4 and 3.2.2.3.5. It is not an inherent property of segments, but is derived from the interaction between segments. It makes different kinds of predications and is not something that could allow us to set up segmental strength scales. The strength that can be gained by a segment in this way through the sharing of individual elements becomes a property of the whole segment and thus accounts for the fact that lenition is inhibited from affecting the segment.

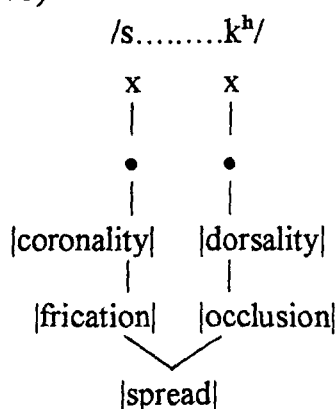
⁹⁹ It may be that, for full geminates, these are best thought of as two ‘half-segments’, but the generalisation is the same.

The idea that element sharing can give strength also accounts for the situation in Liverpool English, where /t^h/ is inhibitory, but /k^h/ was not. It further accounts for much of the inhibition in the HGSC₁, where, for example /ŋk^h/ was shown to be the single most inhibitory environment - it was the only inhibitory environment in High Alemannic in Keller's (1977) 'context 3', as was shown in (2.13). Certain other aspects of the melodic inhibition in the HGSC₁ also follow, for example, the fact that in Mosel Franconian and Rhine Franconian, environment D [c__] was only inhibitory (for /k^h/ and /p^h/) where the consonant was a nasal or in a geminate. It also, naturally, accounts from the partial geminates for the environment [N__] in Bantu, Kannada, Irish, Proto-Iranian and Liverpool English.

Moreover, once this possible avenue of explanation is opened up, other cases of melodic inhibition become comprehensible. We have seen that [s__] is an inhibitory environment in the GCS₁, HGCS₁ and Liverpool English and we might wonder whether this can be accounted for in the same way. Indeed, this fits well with an aspect of Iverson & Salmons's (1995) original proposal which formed part of the basis for the laryngeal realist analysis of laryngeal specifications originally developed in section 3.1.3.1.

Iverson & Salmons (1995) propose that clusters of the alveolar fricative directly followed by /p^h, t^h, k^h/ share a |spread| specification. They take this as the reason why such clusters lack appreciable phonetic aspiration in the stop release phase. Their point is that, given that there is only one |spread| specification, which is shared autosegmentally between the fricative and the following stop, the glottal spreading gesture involved occurs only once (as is indeed to be expected). This means that, by the time that the stop cluster is finished phonetically, the glottal gesture is already over and the phonetic effect of this is that there is little or no aspiration. This is not the same as saying that the stops here are 'unaspirated' (ie, underlyingly non-|spread|), rather, the claim is that they share their |spread| element with the preceding fricative, so while a phonetic representation might be [sk], for example, underlyingly, the cluster is /sk^h/. A representation for this is given in (3.72), for one of the clusters which involves the velar stop (the same basic situation also applies for the coronal and labial stop).

(3.72)



If element sharing is understood to give a cluster strength, representations such as that in (3.72) can be seen to provide an account for why it is an inhibitory environment. It seems that, along with partial place geminates, such as /ŋkʰ/ and /ld/ and the others discussed above, there can also be partial laryngeal geminates, such as /skʰ/. These share their laryngeal elements and it is this that gives the segments involved the same type of ‘strength’ that was outlined above for partial place geminates.

The phonological entities described here which are bound together by autosegmental sharing are, in fact, the phonological ‘molecules’ which we metaphorically predicted in section 3.1.6.2. To maintain the metaphorical approach, we might say that the bonds between the segments involved in the partial and full geminates hold them tightly together, rather as the electron bonds do to the atoms which are found in molecules. When lenition processes are introduced into the phonology of a language, the existence of these melodic molecules can inhibit the lenition in the stop involved. Note that segments involved in these ‘molecules’ retain their identity as examples of /kʰ/, /d/, etc, just as the atoms of oxygen do, for example, in water, but they are also part of a bigger unit.

This approach avoids the problem which was identified by Elmedlaoui (1993) for previous approaches to inalterability because it places the inalterability effects where they belong, as a property of the segments concerned in certain environments, not as a condition on the applicability of rules. A process which only affects geminates, or partial geminates (such as is perfectly conceivable on Hayes’s 1986 approach, but seems extremely unlikely) becomes impossible here. Indeed, it is predicted that such processes will not exist. We return to this point in chapter 5. The approach adopted here also widens the scope of the ‘inalterability’ generalisations by the recognition of partial laryngeal geminates.

The inhibitory effects addressed in this section are melodic in the sense that, for a lenition to be inhibited, the segment which would be the input to a lenition process can only be 'protected' if it is in the right melodic environment. It is to be expected that this melodic kind of effect will interact with other factors, such as the prosodically inhibiting factors which were identified in the last section. This seems to be the case in much of the lenition data that we have investigated here. If we return to the wider set of possible environments from (3.69), we can see how the two types of factors can combine to inhibit the innovation of a lenition process. The least inhibitory environments are prosodically weak and have no melodic support; these are environments A, B and C1, and lenition seems to be common here. Intermediate are those environments which are prosodically strong but have no melodic support; these are environments E and C2, which can indeed be seen to effect inhibition at times. The most inhibitory environments are those discussed in connection with melodic inhibition in environment D (these can often be seen as having prosodic support, too, because the segments occur in an onset, which could be seen as an 'initial' prosodic position); it is in these environments, that no lenition at all is often observed.

3.2.3.3 Summary and prospect

In discussing the environmental factors which are connected with the lenition processes that we have investigated in this thesis, both those which are synchronically active in present-day languages and those which are now fossilised, we have seen that any account of inhibiting environments for lenition cannot hope to apply in the same way to all lenitions. While we can identify both prosodic and melodic inhibiting environments and can generalise over them, they are not the same for all lenitions. Occasionally lenitions occur across the board, without any inhibition and affecting all occurrences of a segment (as in the Greek case from section 3.2.1.6) but it is typically the case that some pattern of inhibition is exhibited.

We have seen that certain clear generalisations can be made for both prosodic and melodic inhibition, but it is also clear that we need to recognise the role that chance can play. It is not the case that all of the potentially inhibitory factors discussed above succeed in inhibiting lenition in every case. This is evident for the prosodic factors: we have seen that environment C1 [\acute{v} __ (v)] is inhibitory in some lenitions (in Liverpool English somewhat and elsewhere) but it is not inhibitory in the Spanish lenitions. Also, E [#__] is inhibitory for Liverpool English and the HGCS₁ because it only allows affrication, not spirantisation, but in the Spanish and Irish cases, it is not inhibitory when it represents [v #__] and it is not inhibitory at all in the GCS₁.

The same is true for the patterns of melodic inhibition. The environment [N__], which, as we have seen, is one of the most robust cases of melodic inhibition, is not inhibitory in the GCS₁, for example (as can be seen by comparing, for example, Gothic *munþs* [munθs] ‘mouth’ with Latin *mentum* [mentum] ‘chin’). This could be interpreted to show either that (i) partial laryngeal geminates are the ‘strongest’ environment, because there are inhibitory in the GCS₁, or (ii) that chance can determine which has an effect in any given language. There are also certain other melodic effects in some of the lenitions that we have not investigated in detail in this section, as can be seen in the discussion of the data. It could be that these indicate the need for future research on the issue or it could be that they represent truly chance exceptions. I leave this issue open here, but return to it briefly in chapter 5, where I show that this does not rob the idea of predictive (and hence potentially explanatory) power.

The firm conclusions from this section are that the initial positions in prosodic domains can function inhibitorily and that autosegmental elemental sharing of the types discussed in the last section can provide a basis for melodic inhibition. It is possible that these conclusions raise as many questions as the answer,¹⁰⁰ but I believe that the questions are worth exploring and that the approach taken here can be seen to help relieve us from puzzlement about the phenomenon, which was the minimal aim adopted in section 1.2.1.

In this wider section (3.2), I have considered a range of issues which are connected with the study and understanding of lenition. The Germanic data from chapter 2 was compared with a number of other lenitions from unconnected languages and drawn out the implications of this. We have seen that ‘lenition’ is not formally unifiable as a process type, but that there are nonetheless commonalities among the processes which allow us to justify dealing with the processes together. As has been shown in here, one of the points that many of the processes have in common is that they can be inhibited by the range of factors discussed in this section. In the next chapter we use the knowledge and understanding of the issues that we have gained here to reanalyse in detail the Germanic data which was first introduced in chapter 2.

¹⁰⁰ Some particularly interesting open questions are: (i) what is the true ‘cause’ of prosodic inhibition?, (ii) what is the precise mechanism through which elemental sharing can inhibit lenition?, (iii) are there cases of interaction between vowels and consonants in terms of elemental sharing, which can also be shown to give ‘strength’? In terms of (iii), I think the answer is yes, but it requires us to reinterpret some of the elements used here to tie them in with the elements found in vowels. It would take us too far afield to discuss this in detail, as argued in section 3.1.3, but if [labiality] is one way of interpreting [U], as has often been argued, then this can provide an explanation for why the vocalic change /u/ > /ʌ/, which was innovated in certain varieties of Early Modern English, was often inhibited when the original /u/ was next to a labial consonant (see, for example, Dobson 1968). We could interpret this as strength given by the sharing of [U] between the vowel and a neighbouring consonant. This is the kind of example of which more need to be found to back up the idea. I believe the signs are promising, but I leave this to future research.

4 Historical and theoretical phonology: lenition in Germanic languages

Throughout this thesis, it has been overtly acknowledged that one of the main aims of the discussion is to enable an informed understanding of some specific historical data. We encountered this data first in chapter 2, where it was set out in some detail, relying on the achievements of philological investigation into the history of the individual languages concerned and resting on evidence provided from the comparative method of phonological reconstruction, from the interpretation of spelling in texts and from the study of present-day dialectology and the ways in which we can project that back to earlier linguistic stages.¹ The data was split there into four sets, each of which was given a name, and this is the way in which individual pieces of the data have been referred to during the discussion in chapter 3, where we saw how some of the data can be fruitfully viewed as examples of general phonological process-types, especially when compared with analogous processes from other languages. When viewed in this way, the data can play an important role in informing phonological theory about the nature of possible phonological processes and the kind of patterning that they can be seen to involve. As we have seen, the use of such historical data can make a substantial contribution to our general understanding of phonology (in chapter 3 this included helping to provide the foundation for the 'laryngeal realist' view of laryngeal phonology, for an understanding of what we might really mean when we speak of 'lenition processes', and for the recognition of the facts of lenition inhibition and the formulation of a possible theoretical means of understanding it). In this way, historical phonology can have implications for theoretical phonology. If diachronic phonological events involve the innovation of novel processes into the phonology of a language, then we can expect the constraints that this phonology puts on such process innovation to be enlightening about its own deeper nature.

Theoretical phonology also has substantial implications for the study of the historical phonology of individual languages and language varieties, as well, of course, and I explore some of these in this chapter. In chapter 3, as has just been explained, the Germanic processes were viewed simply as examples of processes, in a universalist light, but they are also interesting in their own right as occurrences in the history of the languages that they were innovated into. I consider them individually in the chapter. This will involve a much more detailed consideration of them than they received at various points in chapter 3. The essentials of the analyses that were developed in that chapter for the types of processes that they involve

¹ This chapter seeks to contribute in a small way to the philological interpretation of some of this data. In particular, certain aspects of the orthographic data given for the processes in chapter 2 will be reinterpreted.

will not change here, but returning to and reanalysing the Germanic developments in the light of our increased understanding of phonology in general and of lenition in particular will not be a simplistic or pointless task. As will become apparent, the careful consideration of these diachronic occurrences in their own right and in their linguistic context will lead to a sometimes substantial revision of our understanding of it. In other cases, while the phonological analysis of a process involved here itself may be straightforward, the implications of its innovation into a particular system can still be intriguing.

The 'sets' of data established in chapter 2 are returned to as groups here. It will be clear from the discussion which occurred already partly in chapter 2 and was reinforced in section 3.1.6 that this decision is fundamentally done in respect for philological tradition. There is no reason to believe that the 3 'parts' of the GCS, for example, were causally linked to each other. In the discussion in this chapter, I therefore deal with each of the processes involved in these groups separately, in three of the subsections to section 4.1.1. A final subsection, 4.1.1.4, considers certain points which arise from considering all of the processes together, however. The other three sets of data are considered individually in sections 4.1.2 to 4.1.4. There is also a short second part to this chapter (section 4.2). This considers certain aspects of the laryngeal history of two present-day Germanic languages, 'standard' German and Dutch. The relevance of the discussion there will become apparent in those sections; at least certain aspects of the points discussed there have only been rendered moot with the recognition of laryngeal realism.

The discussion in this chapter will sometimes involve a brief consideration of previous proposals for the analysis of the processes, particularly when such analyses exist from the literature in the tradition of theoretical historical phonology. As was remarked in chapter 2, there exists a wealth of such analyses for some of the data covered here (especially for the GCS and the HGCS), whereas there are far fewer for the others (the EIFV and especially the IGCW). Very few such analyses can be considered here however and it will be clear by now that I consider that much previous work has missed certain important details of analysis and important generalisations in possible patterning, so I focus on what contributions I believe such work can make to our understanding of the data by considering it in the light of the general model of phonology and the specific way of understanding lenition which were developed in chapter 3.

We left chapter 2 with the data somewhat reanalysed from its standard philological presentation. The discussion in section 2.3 considered several important points which are relevant to the phonological interpretation of the philological data. When they were

considered at various points in chapter 3, certain aspects of the Germanic data were slightly reanalysed, most particularly in terms of laryngeal realism. When I turn to each of the processes concerned here, I take this into account. Some of the analysis in chapter 3 has already explained all that there is to say about some of the data from chapter 2, however, specifically in the case of the so-called ‘GCS’ debuccalisation. I do not discuss this particular process any further, therefore, but rather focus on those processes where there is still something to say.

4.1 Germanic lenitions and laryngeal specifications

The discussion of laryngeal specifications in section 3.1.3.1 was founded on the idea that languages which have two series of obstruents can be divided into |spread| languages and |voice| languages (assuming that the language is not a |constricted| language, of course), and that the distinction rests on phonetic interpretation and phonological argumentation. The discussion there showed that most present-day Germanic languages are |spread| languages, (or, rather, that most varieties of most present-day Germanic languages are |spread| languages). The discussion there also included a section which focussed on the reconstruction of the situation in Proto-Germanic (section 3.1.3.1.10). We saw there that both comparative and phonological evidence point unambiguously to the fact that Proto-Germanic was a |spread| language.

This is naturally highly important for the discussion here. It gives us a ‘middle point’ for the processes that are to be analysed because the Proto-Germanic phonological system was derived from that of Indo-European through the GCS (and other phonological processes) and we can assume that the segments of Proto-Germanic formed the input for the HGCS, EIFV and IGCW because the languages in which these processes occurred are all derived from Proto-Germanic (no doubt through various stages which we might label West Germanic, North-Sea Germanic, etc).

The Proto-Germanic stop system was therefore /p^h, t^h, k^h/ : /p^o, t^o, k^o/. In this chapter we will need to consider what the consonantism of Indo-European was, because, of course, this is linked to the Germanic system through the GCS. In our contemplation of the GCS here, we will need to consider what type and number of processes were needed to map IE onto Germanic. This will not be helped by the uncertain status of the Glottalic theory in IE reconstruction (a fact illustrated by the two options in the ‘take 4’ version of the GCS which

was the last one considered in chapter 2). The picture presented here will naturally be geared towards a consideration of what makes sense for Germanic. While it would be wrong to ignore the developments from IE into other daughter languages, it would take us too far from the focus of discussion here to consider them all in detail.

Apart from the relationship between the GCS and the other processes, we will see that certain other relationships can be made out between the processes which are to be discussed here. Now that we are armed with an understanding of phonological structure, phonological processes, process innovation and what we mean by 'lenition', it is indeed to be hoped that this can all be profitably applied to the data to be discussed.

4.1 The Germanic Consonant Shift revisited

The discussion in chapter 2 showed that there are three parts to the GCS. This was noted in section 2.3.2, when we recognised the effect that the Laryngeal Theory has had on the reconstruction of the consonantism of Indo-European. As far as Germanic is concerned, there was, in fact, never any evidence for a fourth IE stop series of *TAs*. We have also seen that the mechanism by which it has been proposed (initially by Kuryłowicz 1927) that the *TAs* were formed in Indo-Aryan (ie, 'fusion', as discussed in section 3.1.3.1.8) is backed up by phonological theory, and has analogues in a range of other languages, so it seems beyond doubt that there were three series of stops in IE. The precise laryngeal specifications of those series is not so clear, however. At the end of chapter 2, we saw that there is a fair amount of evidence which speaks for the Glottalic Theory of IE stops, but that there are also problems with the idea, not least of which is that there are hardly any ejectives in present-day IE languages. The only case where there are ejectives is in certain dialects of Armenian (see Gamkrelidze & Ivanov 1984) and these are generally thought to have been borrowed from neighbouring non-IE languages which make substantial use of such [constricted] stops.² This type of linguistic borrowing, which involves borrowing the basis of a segmental contrast is far from unknown; as was discussed in section 1.3, we can be sure that it has occurred in the areal spread of clicks in southern Africa, for example. The occurrence of ejectives in Armenian cannot be taken as evidence for the Glottalic theory, therefore. It is also not necessarily the case, however, that the general lack of ejectives in IE rules out the Glottalic Theory immediately. The factors which led to the proposal of the idea cannot be entirely

² Even Gamkrelidze & Ivanov (1984) who are among the most enthusiastic proponents of the Glottalic Theory entertain the idea that these ejectives are not inherited directly from a 'Glottalic' IE.

ignored, and the assumption of the Glottalic Theory does allow for a typologically reasonable (or maybe 'phonologically possible') reconstruction of IE and for principled answers as to why there were certain co-occurrence restrictions among consonants and why there were (probably) no occurrences of the labial stop which was traditionally transcribed as /b/.

It has been contended (in, for example, several of the papers in Vennemann 1989a) that the Glottalic Theory faces a serious problem because it assumes that the ejectives became voiced stops in most of the IE daughter languages (such as Italic and Slavic). Critics propose that this is not a possible phonological innovation and so should not be assumed for IE, but recently Fallon (1998) has shown that several cases of ejective voicing are attested in various languages, so this criticism is not devastating to the Glottalic Theory.

In this section, I consider both a 'Glottalic' option and a non-Glottalic option for the GCS processes. I discuss the processes individually in sections 4.1.1.1 to 4.1.1.3, given the fact that there is no obvious causal relationship between the different parts,³ and then return briefly to consider them in the context of each other and of 'Verner's Law', which, as is well known, has some relationship with GCS₁. As a starting point, I illustrate the latest formulations of the GCS as (4.1). This is a 'take 5', rather than the 'take 4' which was the last one encountered because it has been minimally reformulated from (2.37) to illustrate the laryngeal realist interpretation of aspiration (for the inputs, but not the outputs),⁴ as argued for in section 3.1.3.1; this is compatible with the general 'elemental' approach to phonology developed in section 3.1. This interprets a situation where there is 'allophonic aspiration' (which Gamkrelidze & Ivanov 1973 assume for the IE *T*s and *MA*s) as an indication that the segments are 'underlyingly aspirated' (ie, characterised by |spread|), with a loss of the element in 'non-aspirated' environments. This means that the brackets around the symbol for aspiration in the *T*s and *MA*s can be removed. One of the consequences of this is that the input

³ In discussing the options of whether the processes of the GCS were independent or interdependent, Collinge (1985) writes that "[i]ndependence seems never to have won serious adherents" (1985, 66) and describes a range of 'chain shift' analyses. However, it still does not seem necessary to me to posit a connection between the processes simply because they must have occurred within the period of several hundred years of each other. The analyses of, for example, Prokosch (1939), Forquet (1948) and Davenport & Staun (1983) seem ingenious but unnecessary. Prokosch seeks to link the processes as all being an expression of some abstract change in the constriction in the vocal cords, Forquet sees them all as a quite large number of processes which each moved all three series on by one step at the same time, and Davenport & Staun present a Dependency Phonology analysis in which the individual processes are seen as steps in the shifting of |V|s and |C|s towards an ideal Germanic obstruent system, but it is not clear what mechanism could be thought to recognise the need for this in a language and hence institute the processes. Given the understanding of the innovation of lenition processes which was developed in section 3.2.2.3, it will be clear that such accounts are problematic because they do not express the phonetic or phonological reasoning which can be linked with the explanation of the processes.

⁴ I deal with the outputs in the individual subsections below.

for the GCS₂ is seen as the same in both the traditional account and in (Gamkrelidze & Ivanov's) Glottalic Theory. The formulations of the GCS can thus be represented as:

(4.1) GCS: **take 5** ('traditional' or 'Glottalic')

GCS₁

$$\begin{aligned}\{p \text{ or } p^h\} &> f \\ \{t \text{ or } t^h\} &> \theta \\ \{k \text{ or } k^h\} &> x \\ \{k^w \text{ or } k^{wh}\} &> x^w\end{aligned}$$

GCS₂

$$\begin{aligned}b^h &> \beta \\ d^h &> \delta \\ g^h &> \gamma \\ g^{wh} &> \gamma^w\end{aligned}$$

GCS₃

$$\begin{aligned}\{b \text{ or } p'\} &> p \\ \{d \text{ or } t'\} &> t \\ \{g \text{ or } k'\} &> k \\ \{g^w \text{ or } k^{w'}\} &> k^w\end{aligned}$$

One of the positive things that has resulted from the proposal of the Glottalic Theory is that the debate has been opened up about what is the best formulation of the correspondences between IE and Proto-Germanic (for example) and the processes which relate them. The formulations that I propose here involve a certain reinterpretation of the symbolic units which are used to represent them. In chapter 2, some recent work by Vennemann (eg, 1984, 1994) was discussed which casts doubt on the traditional interpretation of the GCS and its connection with the HGCS. Vennemann proposes a 'Bifurcation Theory' of Germanic developments which focuses particularly on what is described here as the GCS₃. I address this below in those sections where it is relevant.

4.1.1.1 GCS₁

The GCS₁ is probably the best known, most secure and most discussed 'part' of the GCS. Collinge (1985) notes that it is the part which is most unscathed by the Glottalic Theory, in that all formulations of IE consonantism assume that a spirantisation has occurred. In

Vennemann's (1984, 1992) Bifurcation Theory, the GCS₁ is singled out as the only part which is described as a common Germanic development (the others, he claims, occurred differently in different dialects). Vennemann (1984) labels this part alone as the 'Proto-Germanic sound shift'. We have already seen in section 3.2.2.3.1, the segmental analysis that will be adopted here, and in this section, I make this explicit and consider some of the implications.

The GCS₁ is clearly a lenition under the definition developed in section 3.2.2.3. It cannot be attributed in a causal way to its environment, so it is weakly unconditioned. It could not possibly be seen as assimilatory, for example, as it occurred in all prosodic environments, including the word-initial environment, which is often inhibitory. We saw in section 3.2.2.3.1 that true spirantisation of this type (ie, that which actually has fricatives as its output) cannot be achieved through the loss of elemental material. Rather, it was proposed there that spirantisation can only naturally be innovated through the interaction with a previous affrication, which is linked to a |spread| specification. The conclusion from this is that the original input for the GCS₁ must have been a |spread| voiceless stop and that the correspondence between stop and fricative actually involves two quanta, and initial affrication, followed by a subsequent 'deaffrication' (or 'true spirantisation'). The original affrication introduced a |frication| element through the acoustic reanalysis of the effect of the |spread| element (ie, aspiration). This realisation in fact provides for a link with a long tradition of analysis of the GCS₁ by overtly tying it in with the presence of aspiration in the stop inputs, although the details of the analysis differ somewhat. As Schrodtt (1976) notes, "[a]gain and again, a series of *TAs* has been assumed as an intermediate stage in the shift of the *Ts*" (1976, 218).⁵ Previous approaches have not sought to account for the mechanics of the development from /p^h, t^h, k^h/ to the fricatives, however. The approach developed in section 3.2.2.3.1 does this, by showing that an intermediate affricate stage can allow for an analysis which is compatible with the natural and minimalist approach adopted in this thesis.

The approach adopted here seems to be compatible with Gamkrelidze & Ivanov's (1973) reconstruction for the *Ts* of IE, which sees them as being |spread| stops. However, there is a problem with this latter assumption. This is the fact that, as we saw in sections 2.3.2 and 3.1.3.1.9, the *Ts* of IE were turned into |spread| stops in Indo-Iranian through a process of fusion with following laryngeals, so it cannot have been the case that Indo-European had a

⁵ The original reads "[i]mmer wieder wird als Zwischenstadium der T-Verschiebung eine Reihe von *TA* angenommen." Iverson & Salmons (1995) also cite this passage and note the connection between this tradition and their approach to laryngeal specifications (which is largely taken over in this thesis).

series of such stops to begin with. I propose that in Pre-Germanic (that is, late-Indo-European, during its break-up into dialects), exactly the same process of fusion occurred, so that the IE *T*s which were adjacent to IE laryngeals picked up |spread| to become ‘voiceless aspirated’ stops, as in Indo-Iranian. Unlike in Indo-Iranian, however, this process was generalised throughout the Pre-Germanic *T*s so that all occurrences of the segments involved became |spread| stops. In what follows, I assume this stage of late-IE/Pre-Germanic as the starting point for the GCS₁.⁶

I present a final analysis (‘take 6’) for the GCS₁ in (4.2) which illustrates the two processes involved from the late-IE starting point. In line with the approach adopted in section 3.2.1, following the justification in section 3.1.3.1.11, there is no need to assume any change in the laryngeal specifications of the segments involved (indeed, given the assumption of minimal quanta, this would require the proposal of a separate process).

(4.2) GCS₁: final version

$$p^h > pf^h > f^h$$

$$t^h > t\theta^h > \theta^h$$

$$k^h > kx^h > x^h$$

$$k^{wh} > kx^{wh} > x^{wh}$$

As we saw in section 3.2.3, the first of these processes was inhibited slightly. There was melodic inhibition in partial laryngeal geminates and a small number of other melodic cases which seem more obscure (lenition of the alveolar was inhibited following an IE labial or velar *T*, which did spirantise itself).⁷ It is possible that these other inhibitory environments also represent partial geminates, where |spread| was shared, and could provide inhibitory strength (this is what Iverson & Salmons 1995 assume, although they do not formulate it as a laryngeal geminate effect) or it may simply be that this is an inexplicable case of chance inhibition. In any case, the lenitions here were remarkably uninhibited, with no prosodic inhibition at all, just as in the Greek case which was discussed in section 3.2.1.6.

⁶ This means that at least this part of Gamkrelidze & Ivanov’s reconstruction cannot be correct, on the laryngeal realist interpretation given to it here. This does not have any effect on the other aspect of their proposals, however.

⁷ Note that the generalisation cannot be captured as a general Germanic prohibition of a sequence of two fricatives. in Proto-Germanic. As Mouton (1972, 1954) clearly shows, such sequences did exist in Proto-Germanic, as among others, Davenport & Staun (1983) note (for example /xs/ in words which are inherited into German as *sechs* ‘six’, *Lachs* ‘salmon’).

4.1.1.2 GCS₂

On the formulation given in (4.1) in the last section, the GCS₂ seems to be relatively straightforward. While the discussion in section 3.2.2.3 will have clear implications for the processes involved, it seems to be a straightforward unambiguous case of lenition. As we will see in this section, however, certain aspects of the traditional philological description of the GCS₂ are problematic and some of the reasoning and observations in section 3.2.3 will have important implications for how we understand the processes. Some of the points made here have been discussed before, but they have not been put on a firm foundation.

Much of the interest here will derive from the patterns of inhibition which are reconstructed for the GCS₂. In section 2.1.1, the GCS₂ was described as being strongly unconditioned, that is, as being entirely uninhibited. We will see below that there is neither reason nor need to assume this, and indeed, that there is every reason to believe something quite different. I turn to these points directly, but there are a few other aspects of the GCS₂ which have been shown to be problematic, and I deal with these first. If we consider the formulation of the processes in (4.1), we can see that the symbols used for the output are those for fricatives. We saw in section 3.2.2.3.2, however, that spirantisations without *|spread|* do not lead to fricatives. Rather, this type of process is typically approximantisation. We are thus brought to assume that the GCS₂ is in fact not a true spirantisation, but should be better represented as in ‘take 6’ in (4.3), with approximants as outputs.

(4.3) GCS₂: take 6

$$\begin{aligned}b^h &> \beta \\d^h &> \delta \\g^h &> \gamma \\g^{wh} &> \gamma^w\end{aligned}$$

The other point of interest, before we turn to lenition inhibition, is the precise nature of the laryngeal specifications involved in the segments. We saw in section 3.1.3.1 that both Lombardi (1991) and Iverson & Salmons (1995) analyse ‘voiced aspirated’ stops (such as those of Present-Day Hindi and those which are reconstructed here as *MA*s for IE) as containing both *|voice|* and *|spread|*, and Harris (1994) proposes the same, as well.⁸ A subsegmental representation for such segments was given in section 3.1.5 as (3.31). We can assume that the *|spread|* specification was lost between IE and Proto-Germanic, because

⁸ Of course, none of these analysts use the precise terminology that I have adopted here, but when it is translated into the terms of their frameworks, the analysis is the same.

voiced aspirated segments are not attested in any of the Germanic languages, and this is represented in (4.3). A fair question here is whether this loss of |spread| occurred in the same step as the approximantisation, as is actually shown in (4.3). We return to this point in this section below, where we will see that the question is not quite the right one to ask. We can also note here that on this formulation, the fricatives are left with a |voice| element. We also return to the fate of this element below. For the moment, we can note the parallel between the inputs and outputs of the GCS₂ and those of the Spanish approximantisation, which results from their analogous origin through the loss of |occlusion|.

The picture of the GCS₂ in (4.3) is still problematic, however. The problem comes to light when we consider the GCS₂ in the context of the full history of the languages into which it was innovated. If we consider the data that was taken from the Handbooks and offered as evidence for the processes in section 2.1.1, the problem becomes apparent. The data is repeated here from (2.3) as (4.4).

(4.4)

b ^h	Old Indic	<i>nábhas</i>	Old Saxon	<i>nebal</i>	‘mist, cloud’
d ^h	Old Indic	<i>rudhirás</i>	Old Norse	<i>rauðr</i>	‘red’
g ^h	Latin	<i>hostis</i>	Gothic	<i>gasts</i>	‘enemy, foreigner’
g ^{wh}	Greek	ομφη	Old Norse	<i>syngva</i>	‘voice, sing’

It is clear from the data here that at least some of the segments which are actually attested in the Germanic languages are stops. This is true for both the spelling in the earliest documents, as in the examples for the Gothic and Old Norse reflexes of IE /g^h/ and /g^{wh}/ in (4.4), and in the Present-Day Germanic languages. The Latin and Greek cognates certainly show their own effects of phonological change from IE, in fact, it seems that the Germanic words *gasts* and *syngva* show evidence for much less change. There are indeed some words in Germanic which have fricative reflexes for the IE *MA*s (Old Saxon *nebal* and Old Norse *rauðr* are fair examples). But there are also many others which show stop reflexes. Some more of these are given in (4.5), taken from Krahe (1969) and Luick (1914-1940).

(4.5)

b ^h	Old Indic	<i>bahávati</i>	Gothic	<i>bauan</i>	‘is, live’
d ^h	Old Indic	<i>bandhati</i>	Old Frisian	<i>binda</i>	‘bind’
b ^h	Old Indic	<i>abhi</i>	Old English	<i>ymbe</i>	‘around’
g ^h	Latin	<i>homo</i>	Old English	<i>guma</i>	‘man’

The standard way of accounting for this, in line with the picture that has been presented up till now for the GCS₂, is to assume that there was a process of occlusivisation following the GCS₂ at some point in the history of Germanic. Lass & Anderson (1975) formalise this as a phonological rule, for example, and Prokosch (1933) even names it ‘the intermediate shift’. On such an analysis, it must be assumed that this occlusivisation process was a common Germanic development, because most of the reflexes of the IE *MAs* are stops in all Germanic languages.

We have not encountered occlusivisation previously in this thesis, and this is at least partly due to the fact that such processes are very rarely reported in the history of languages. Occlusivisation is the inverse of approximantisation in that the outputs and inputs are the opposite way round, and this would make it a case of ‘fortition’. In section 1.4, the concept of fortition was very briefly mentioned. I explained there that I do not consider it in any detail in this thesis because it has a very questionable place in phonology. I claimed that there are few unambiguous cases of the endogenous innovation of fortition processes in languages. The ‘intermediate shift’ occlusivisation process described here, though, seems to be a substantial counterexample to that claim. However, there is, in fact, no real reason to assume that this aspect of the history of Germanic languages requires us to recognise both an approximantisation and an occlusivisation.

In order to understand precisely what occurred, it will be helpful to consider what was the synchronic situation in Proto-Germanic. Once the labiovelar was lost,⁹ as Moulton (1972) shows after an extremely detailed consideration of the evidence, there was allophony among the three underlying segments which were the outputs of the GCS₂. Moulton’s (1972) summary is given in (4.6); as this is a direct quotation, I retain his original ‘philological’ symbols:

(4.6)

- ... /b ~ ʙ/. All the oldest Germanic languages agree in showing the stop [b] initially, in gemination, and after a nasal...
- ... /d ~ ɖ/. All the oldest Germanic languages agree in showing the stop [d] initially, in gemination, and after a nasal, and after /l/, /z/ and /g/...
- ... /g ~ ɡ/. All the oldest Germanic languages agree in showing the stop [g] in gemination, and after a nasal...

Moulton (1972, 173)

⁹ There is actually no evidence for a labiovelar in this series in Germanic. While Gothic spelling shows that both /x^{wh}/ and /k^{wh}/ existed in Germanic, there is no evidence for distinct reflexes of IE /g^{wh}/. These (or the input segment itself) were probably lost early in Germanic.

In line with the traditional position, Moulton concludes that “PIE [bh dh gh] ... seem first to have given pre-Germanic. [b̥, d̥, g̥] in all positions. But then-gradually, and in more and more positions ... [b̥, d̥, g̥] became [b d g]” (1972, 172). It seems to me that this conclusion is unnecessary, however, as intimated above. I deal with some of the reasons why it has been assumed that the GCS₂ was an entirely uninhibited lenition below, but first I note some arguments against the idea.

Conceptually, it seems problematic to assume that the segments first underwent an approximantisation and then underwent an occlusivisation which precisely reversed the results of the approximantisation. If we put aside the loss of [spread], then the processes which would be involved are basically formalisable (in an informal rule-based format) as in (4.7).¹⁰

- (4.7)
- a) b d g > β̥ ɸ̥ ɰ̥
 - b) β̥ ɸ̥ ɰ̥ > b d g / $\left\{ \begin{array}{l} \# __ \\ \text{in geminates} \\ \text{in partial geminates} \end{array} \right.$

The two processes are practically a mirror-image of each other and together form a diachronic ‘Duke of York’ analysis.¹¹ Given the oddness of a history which reverses on itself, an analysis which does away with this in favour of a simpler, more natural history would seem conceptually more likely. Given the understanding of the patterning of the lenition inhibition which was developed in section 3.2.3, the traditional analysis seems highly unlikely, because it proposes that no inhibition occurred in the lenition (4.7a) but then that occlusivisation occurred in precisely those environments which we have seen to be inhibitory in lenitions. An alternative analysis seems much more likely than the proposal that there were these two, chronologically distinct processes, one of which had to undo the effects of the other. The simpler proposal for this aspect of the GCS₂, is that there was simply one general lenition, which was inhibited both prosodically and melodically in ways which are very common, as we saw in section 3.2.3. If we consider the environments where Moulton (1972) recognises stops, as described in (4.6), we can see that prosodic inhibition would account for the stop reflexes in word-initial position and melodic inhibition accounts for the stop reflexes in

¹⁰ There are a few caveats that need to be made about the form of the ‘rules’, and these will be recognised below.

¹¹ Pullum (1976) introduced the idea of the ‘Duke of York’ derivation into linguistics. He uses the term to describe lengthy synchronic phonological derivations which work with ordered rules and could synchronically change segments into something during a derivation and then change them back again into what they were before.

geminate, the [N__] partial place geminate, and for /ld/ and probably for /zd/ as well. The only slight surprise is /gd/, which certainly invites further consideration, but which could simply be a chance inhibitory effect. Given that the environments involved here are practically all well recognised as being inhibitory to lenition, it seems more likely that they never let the lenitions occur, rather than that they allowed them, only for the original segments later to be restored in them. On the analysis that I propose here, the lenition basically occurred in intervocalic and final positions, which, as we saw in section 3.2.3 are generally uninhibitory.

In fact, this type of analysis has been proposed before. Vennemann (1984) is clearly of the opinion that the two-process analysis, including occlusivisation, is highly suspect. Indeed, he states that he “would even go so far as to doubt that this is a possible phonological change” (1984, 8).¹² Vennemann describes a distinction between a ‘majority opinion’ and a ‘minority opinion’ on the issue. The majority opinion is the traditional account, which is that given in section 2.1.1 for the GCS₂, and which requires two processes to account for the actually attested forms (or maybe even three processes, if the loss of |spread| was a separate event). Vennemann and some others (this includes the excellent company of Luick 1914-1940 and Strang 1970) have held the minority opinion that the underlying change was from /g^h/ to /g/; this is a position to which I associate myself here.

Of course, this position does not need to claim that there was no approximantisation. It seems right to characterise the GCS₂ as the loss of |spread| in an appreciable diachronic event. The output stops were then subject to an approximantisation which was very similar to that which we saw in Spanish in section 3.2.1.2, and can be assumed to have derived surface approximants from the Proto-Germanic stops in certain environments. The formulation of these environments is entirely in agreement with the generalisations and observations which were made concerning the notion of lenition inhibition in section 3.2.3. This gives us the scenario in (4.8), which represents two synchronic stages. The formulation of (4.8) differs slightly from those given for some of the other processes because it includes information on allophonic realisation. At the second stage, (4.8) claims, both [b, d, g] and [β, ɸ, ɰ] were surface segments.

¹² He writes: “ich möchte sogar bezweifeln, daß es sich hierbei um einen möglichen Lautwandel handelt.”

(4.8) GCS₂: take 7

$$\begin{aligned}b^h &> b \rightarrow \beta \\d^h &> d \rightarrow \delta \\g^h &> g \rightarrow \gamma \\g^{wh} &> g^w \rightarrow \gamma^w\end{aligned}$$

We need to recognise one further aspect of the phonology of Proto-Germanic before we can be satisfied that the full story had been told, however. It was argued in section 3.1.3.1.10 (and noted above in this section) that Proto-Germanic was a |spread| language. This reconstruction holds for a certain stage of Germanic, once the GCS processes have removed the old three-way stop distinction. Now, the output stop segments in (4.48) are the *M*s of Proto-Germanic and these are indicated in (4.8) as containing |voice|. We must conclude that the |voice| element was lost in Germanic as well. There are two ways of conceiving of this: (i) the loss of |voice| was a separately innovated process, or (ii) the loss of |voice| occurred at the same time as the loss of |spread|. Both of these seem possible and the decision as to which is considered to be the more likely really depends on an understanding of quanta. The question here is whether two elements can be lost at once, or whether the maximal quantum is the loss of one. In this case, the data under consideration does not supply an answer, so without evidence to the contrary, I assume that both |spread| and |voice| were lost in the same process.¹³ This seems reasonable given their shared status as laryngeal elements¹⁴ and would truly merit the name ‘delaryngealisation’. This gives us the final version for the GCS₂ in (4.9), which contains two clear lenitions - delaryngealisation and approximantisation.

(4.9) GCS₂: final version

$$\begin{aligned}b^h &> p^o \rightarrow \phi^o \\d^h &> t^o \rightarrow \theta^o \\g^h &> k^o \rightarrow \chi^o \\g^{wh} &> k^{wo} \rightarrow \chi^{wo}\end{aligned}$$

This understanding of the GCS₂ processes allows a natural and minimalist picture to be constructed without the need for overcomplex diachronic derivations or the postulation of a widespread fortition process of a type unattested in other languages. The delaryngealisation

¹³ It is to be hoped that future investigation of the issues may shed light on whether the loss of one always involves the loss of the other. This would involve the study of the diachronic phonology of languages which had such segments. If it is found that the two can be lost separately, then we might reformulate the processes given here to involve a additional stage.

¹⁴ This may cause us to recognise some node-like structure in elemental organisation, which was an option discussed in section 3.1.5. and is a widely adopted solution to such questions (in Feature Geometric frameworks, for example).

follows from the type of IE *MA* which is common to the traditional model of IE consonantism and the laryngeal realist version of the Glottalic Theory. The approximantisation involves the loss of [occlusion] in neutral stops to give neutral approximants which we saw written above as <ɸ, ɔ>.

The approximantisation was subject to dialectal variation, just like some of the other lenitions that we have discussed (the HGCS₁ and the IGCW for example, as explained in chapter 2). The recognition of this final point allows us to account for the data which lead many scholars originally to assume that the GCS₂ involved and was followed by an uninhibited approximantisation and an occlusivisation.

In the 'majority' position, it has been proposed that German exhibits a more widespread occlusivisation than other Germanic languages, because words such as *Liebe* (English cognate: 'love') and *Weib* (English cognate: 'wife') have stops, unlike the other languages, but this, too, is interpretable the other way round. We can simply assume that the approximantisation was further inhibited in those dialects which formed the basis for the varieties of German concerned (which themselves formed the basis for the standard *Standardlautung*). Equally, we can see that Dutch had less inhibition than other varieties; this accounts for the lack of a velar stop in *gaan* [xɑ:n] 'go' and *groen* [xru:n] 'green'. The simple assumption here is that the lenition was not inhibited in those dialects which came to form the basis of the varieties of Dutch concerned (which have also come to form the basis of standard Dutch - *Standaardnederlands* or *Algemeen Beschaafd Nederlands* - see van Bree 1987).¹⁵ In these varieties of Pre-Dutch, the lenition of the underlying velar (and labiovelar) stop was entirely uninhibited, except in partial place geminates of the type [N__] (for example in words such as *koning* 'king' and *zingen* 'sing', as, for example, van Bree 1987) shows.¹⁶ This illustrates again how melodic inhibition can still have an effect in lenition, even when there is no prosodic inhibition.

The analysis of the GCS₂ proposed here in (4.9) allows us to link the diachronic phonology of the Germanic languages to the detailed model of phonology that was developed in section 3.1. It also takes account of the understanding of lenition processes which was developed in section 3.2. The major 'Germanic' lenition is the delaryngealisation and it is possible that the synchronic approximantisation was innovated later, variously in the different dialects. I turn

¹⁵ The possibility is also open here that this lenition of the initial velar occurred later in Dutch, as a separate innovation.

¹⁶ The segments here have been subject to other later processes, of course, including a post-nasal velar loss. I address some of these points briefly in chapter 5.

now to the final ‘part’ of the GCS, before noting how the analyses developed for the separate parts fit together.

4.1.1.3 GCS₃

The GCS₃ involves the one aspect of IE consonantism which is most clearly affected by the Glottalic Theory. One of the main motivating factors behind the proposal that IE had a series of ejective stops, which underwent the GCS₃ in Germanic, is that the labial in the series was quite possibly entirely absent in IE. In the formulations that have been given for the GCS₃ throughout this thesis, I have included a labial, and I continue with this practice below, but it may be that, strictly speaking, the segment should be removed from the discussion. This can easily be done and would not affect anything in the discussion, here.

As far as Germanic is concerned, on the picture built up in this thesis, it makes little difference whether a ‘traditional’ or ‘Glottalic’ model is accepted for IE. There are substantial implications on other parts of the phonology of IE, of course, and the assumption of an ejective series requires the assumption of a process of |voice| acquisition from Proto-Indo-European into most other IE daughter languages, as discussed at the start of section 4.1. As we saw there, Fallon (1998) has provided some possible analogues for this from non-IE languages. Whichever model is adopted, however, one slight complication must be assumed for Germanic, and I discuss this in the current section. The last ‘take’ that we have encountered for the GCS₃ is reproduced here as (4.10):

(4.10) GCS₃: **take 5** (‘traditional’ or ‘Glottalic’)

{b or p'} > p

{d or t'} > t

{g or k'} > k

{g^w or k^{w'}} > k^w

The problem here is relevant to the outputs of the processes, and it is, in fact, the same problem that we have encountered at several other points in this thesis. The question is: what is the true status of the *T*s? If the *T* outputs are the neutral stops /p⁰, t⁰, k⁰, k^{w0}/, then the GCS₃ can be simply modelled as a delaryngealisation, with either |voice| or |constricted| being lost. If the *T* outputs are the |spread| stops /p^h, t^h, k^h, k^{wh}/, then |spread| must be acquired, in addition to |voice| or |constricted| being lost. The latter option is less theoretically appealing, but it seems that this is what actually occurred. We have seen that we can be quite sure that

Proto-Germanic was a |spread| language, and it is these stops which were specified for |spread|. They formed the input to the HGCS₁ and the IGCW, and as we will see in sections 4.2.1 and 4.1.4, the natural assumptions made there back up the reasoning that we have already seen to show that Proto-Germanic had |spread| Ts.

It is possible that the GCS₃ is the result of acoustic reanalysis, and this would seem more likely from an ejective than from a |voice| stop, given that ejectives have a somewhat noisy release phrase. It seems, therefore, that the positions developed in this thesis adds credence to the Glottalic Theory. My suggestion here is not unique in the literature. It is precisely what Vennemann (1984, 1992) proposes, on somewhat different reasoning, for much of Germanic,¹⁷ and I assume this version here, as shown in (4.11):

(4.11) GCS₃: final version

$$\begin{array}{l} p' > p^h \\ t' > t^h \\ k' > k^h \\ k^{w'} > k^{wh} \end{array}$$

It is also possible that the processes here involved two quanta: first the lost of |constricted| and then the acquisition of |spread|. This would be in line with reports of the acquisition of |spread| in other languages, such as in Haider (1985), and would also be entirely compatible with the traditional reconstruction of IE. While it seems that the GCS₃ involved the processes shown in (4.11), I leave the precise mechanics involved open for future research. As was recognised in section 2.1.1.4, the processes involved here, unlike others discussed in this thesis are not, in fact, easily characterisable as lenitions.

Despite the fact that certain aspects of the discussion in this section have been left open, the essentials of the account for the separate processes of the GCS are quite secure. I now move on to consider certain points which arise when they are considered together.

4.1.1.4 Summary and appendix: putting the GCS together (with 'Verner's Law')

The main aim in discussing the processes involved in the GCS in this thesis is to investigate how they can be understood in the context of a wider understanding of lenition (this is partly why the GCS₃ has received less attention than the GCS₁ and GCS₂). The above sections have

¹⁷ To be specific, it is what he proposes for 'Low Germanic', which he proposes as a grouping of all Germanic dialects apart from those of High German, which he describes as 'High Germanic'. I return to his proposal for High Germanic in section 4.2.1.

investigated this in some detail. It is also interesting to consider how the three parts to the GCS can be viewed with respect to each other. I have argued above that they can and should be treated individually but the reason why they can be grouped together as three ‘parts’ to the ‘Germanic Consonant Shift’ is because they occurred at times which cannot be too distant from each other, that is, they all occurred on the way from Indo-European to Proto-Germanic. One important concern in philology, when processes are postulated for approximately the same point in the history of a language, is to consider their relative chronology. Some discussion was devoted to their absolute chronology in section 2.1.1.3, and nothing will be added to that here, but the analyses that have been proposed in this supersection (4.1.1) have certain implications for the relative chronology of the processes, and I address these briefly in the current subsection. I also address one further point which is clearly connected in some sense to the GCS and which, philologically speaking, requires some consideration. This is the process normally known as ‘Verner’s Law’, which interacts in some way with the GCS₁. The discussion of this will of necessity be very brief, but shows, at least, that the concerns addressed in this section are not problematic for the analyses adopted here.

If we consider the final analyses of the parts of the GCS, given above in (4.2), (4.9) and (4.11), certain aspects of their relative chronology become clear. It seems that GCS₁ must have preceded GCS₃, because the final outputs of GCS₃ are the same type of phonological object as the inputs to GCS₁ and, had the GCS₃ preceded the GCS₁, the segments would have merged and the Proto-Germanic correspondences of IE /pʰ, tʰ, kʰ, kʷ/ (or /b, d, g, gʷ/ on the traditional account) would be /fʰ, θʰ, xʰ, xʷʰ/, as well. We must thus assume that the GCS₁ affrication process, at least, had been lexicalised, and was no longer a synchronically active phonological process before the GCS₃ was innovated.¹⁸ In terms of the relationship of these two parts to the GCS₂, less can be said as the output of the GCS₂ were laryngeally neutral segments. However, if the GCS₃ involved two quanta, which was considered possible above, then the GCS₂ must have followed at least this aspect of the GCS₃ because the intermediate stage of the GCS₃ would involve neutral stops. We saw above that the synchronic approximantisation stage of the GCS₂ was slightly different in the phonologies of the separate Germanic languages, so it may well have been active for a long period and would have affected any neutral stops produced by an intermediate stage of the GCS₃, but there seems to

¹⁸ It is, of course, possible that the GCS₁ spirantisation process occurred after the GCS₃, but I assume the contrary here. In fact, the points made about the relative chronology of the parts of the GCS do not hold for this GCS₁ spirantisation.

be no evidence for this. I therefore tentatively assume that the GCS₂ may have followed the GCS₃ chronologically in the same way that the GCS₃ followed the GCS₁.

Given the points just made, I present a summary of the GCS as (4.12), which features the processes listed in a tentative relative chronology. To summarise: the final stages of these processes illustrate the phonological system of Proto-Germanic, which is shown to be a |spread| language, as we have predicted, and to have a series of |spread| fricatives. The Proto-Germanic *Ms* are shown to be neutral stops which are subject to an approximantisation process in certain dialects.¹⁹

(4.12)

GCS₁: final version

$p^h > pf^h > f^h$
 $t^h > t\theta^h > \theta^h$
 $k^h > kx^h > x^h$
 $k^{wh} > kx^{wh} > x^{wh}$

GCS₃: final version (may have involved a stage of /p^o, t^o, k^o, k^{wo}/)

$p' > p^h$
 $t' > t^h$
 $k' > k^h$
 $k^{w'} > k^{wh}$

GCS₂: final version

$b^h > p^o \rightarrow \Phi^o$
 $d^h > t^o \rightarrow \Theta^o$
 $g^h > k^o \rightarrow \mathfrak{X}^o$
 $g^{wh} > k^{wo} \rightarrow \mathfrak{X}^{wo}$

Of the three ‘parts’ to the GCS, only GCS₁ needs to be considered in relation to Verner’s Law (henceforth ‘VL’). VL is a phonological process named after Karl Verner, who first proposed it in 1875,²⁰ and is typically described as a ‘voicing’ process (we have seen in this thesis that such processes need to be carefully considered, hence the scare-quotes) which accounts for why there are two types of reflex in Proto-Germanic for the IE *Ts*. It is typically described as a process which occurred after the GCS₁ and which explains some of the apparent ‘exceptions to the GCS’ by voicing the output fricatives (this account is given in Kurylowicz 1948, Krahe

¹⁹ The phonological system of the variety of late Indo-European assumed here is thus a ‘Glottalic’ one, which coincides essentially with that of Gamkrelidze & Ivanov (1973) and is thus typologically plausible.

²⁰ See Jespersen (1897) for an affectionate account of Verner’s life and his Law.

1969 and Hogg 1992b, for example). On this account, the outputs of VL merged with the non-stop outputs of the GCS₂, once the processes had been lexicalised. Some data which illustrates the VL effects is given in (4.13). This contrasts a word in the first column from a non-Germanic language with a word in a Germanic language in the second column. The words for ‘seven’, ‘father’ show the effects of VL in Germanic, whereas the words for ‘steal’ and ‘brother’ show segments which were not affected by VL. As can be seen from some of the data in (4.13), the placement of accent in IE was a crucial part of the conditioning environment for VL, as the process only occurred where the syllable which bore the word accent did not directly precede the segment concerned. A segment also had to be non-initial for VL to affect it.

(4.13)

Old Indic	<i>saptár</i>	Gothic	<i>sibun</i>	‘seven’
Latin	<i>clepo</i>	Gothic	<i>hlifan</i>	‘steal’
Old Indic	<i>pitár</i>	Gothic	<i>fadar</i>	‘father’
Old Indic	<i>bhrátar-</i>	Gothic	<i>broþar</i>	‘brother’

I do not propose to analyse VL in detail here, but one important point is connected with the proposal for GCS₂ which has been developed here. As explained above, this proposal sees the GCS₂ as essentially a delaryngealisation, accompanied by an approximantisation which was inhibited differently in the various dialects. This means that we do not need to assume an unnatural occlusivisation.

We can be sure that the Gothic examples in (4.13) such as *sibun* featured non-occluded segments (see Marchand 1973, Braune & Ebbinghaus 1981), as they do in Present-Day English. This is in line with the traditional description of VL given above, which sees it as ‘voicing’ the outputs of GCS₁. However, the Present-Day German cognates of such words have stops (eg, *sieben* /zi:p^oən/). On the account developed here, these would always be stops because there was no approximantisation and no occlusivisation. The problem then arises as to how they could be possible inputs for VL on the traditional formulation.

Vennemann (1984) has shown that this is only an apparent problem, however. He proposes to reanalyse VL so that it occurs before the GCS, affecting the late-IE *T*s (which are understood here to have been /p^h, t^h, k^h, k^{wh}/, as they are by Vennemann) so that they merge with the IE *MAs* (those segments which are the input to GCS₂) in the VL environment. They then simply behave as would be expected for such segments and are subject to approximantisation in those dialects where this occurs. I follow Vennemann’s analysis, and

this allows us to recognise that the correspondences which are due to VL are entirely explicable on the picture developed in this thesis.

This recognition concludes the consideration of the GCS, which has been shown to involve processes which can probably only partly be described as lenitions. We turn now to the HGCS, which will be shown to have certain aspects in common with the GCS, as has often been proposed in the literature, but it will also be shown to illustrate a quite different philological point.

4.1.2 The High German Consonant Shift revisited

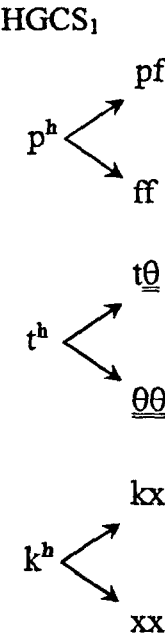
We have returned to the HGCS several times since its treatment in chapter 2. Most of this subsequent discussion has focused on the HGCS₁, rather than the HGCS₂, and we will see in this section that this asymmetry in attention was an informed decision because there is good reason to believe that the HGCS₂ never actually occurred. The HGCS₁, on the contrary, is indisputably a quite substantial process, which, since its lexicalisation, has had a major impact on the underlying segments of those varieties of German where its effects are seen. I discuss certain aspects of two important recent analyses for the HGCS in this section (those of Vennemann 1984, 1992 and of Davis & Iverson 1995 and Davis, Iverson & Salmons 1999) and, as we will see, much of the discussion there focuses on the HGCS₁.

In section 2.3, we began to see how the Liverpool English analogue can help us to understand the HGCS₁. The Liverpool English case has since been discussed in greater detail (in section 3.2.1.8) and we will see here that it can contribute a little more to help guide us to the best analysis of the relationship between affrication and spirantisation, as was explained in section 3.2.2.3.1. Other aspects of the discussion in chapter 3 will also have an impact on the analysis of the HGCS₁; in particular, the account of what is phonologically possible in rhyme structure and in the diachronic behaviour of geminates.

As a starting point for the discussion here, I illustrate the latest formulations that we have seen for the HGCS in the thesis up till now as (4.14). This is 'take 3', rather than simply a repeat of the 'take 2' from (2.32), because it has been formulated to show the laryngeal realist interpretation of the Proto-Germanic segments, but not yet that of the output High German segments. As in the discussion of the GCS in section 4.1.1 above, the nature of the output segments is discussed in detail in this section. I deal with HGCS₁ in subsection 4.1.2.1 and with HGCS₂ in 4.1.2.2. The presentation in (4.14) shows the how the inputs to the HGCS processes are the outputs to some of the GCS processes. The inputs to HGCS₁ are the outputs

of GCS₃ and the inputs to HGCS₂ are the outputs to GCS₂. It is important to remember in this latter case that in the dialects of Germanic which came to form the varieties of German under discussion here, the approximantisation which is given as part of the GCS₂ in its 'final version' in (4.9), did not occur. This is why the inputs to the HGCS₂ are exclusively stops.

(4.14) The HCGS: **take 3**



HGCS₂

p^o > p

t^o > t

k^o > k

The HGCS₁ is formulated in (4.14) to show that, on the traditional account of the HGCS₁, two segments correspond to the Proto-Germanic inputs. This is the first point that we address in the discussion of the processes which we turn to now.

4.1.2.1 HGCS₁

In the first discussion of the HGCS₁, in section 2.1.2.4, it was noted that the processes could be seen in terms of the split that they are presented as in (4.14), or they could be seen as a two stage process which involved first affrication, then spirantisation. This was first presented in (2.16). The discussion of the Liverpool English data in section 3.2.1.8 showed that this analogous, but synchronically observable affrication and spirantisation process is best understood in this two-stage manner because all stages are observable in some environments.

This tied in with the reasoning presented in section 3.2.2.3.1, where an analysis for precisely this type of development was proposed. It seems clear, therefore, that the HGCS₁ should be conceived of as in (4.15), which illustrates the two lenitions as two quanta. Naturally these were inhibited differently, and only some of the original Proto-Germanic *T*s became fricatives.

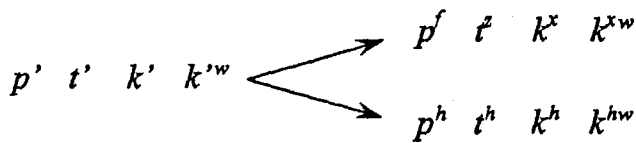
(4.15) HGCS₁: take 4

$$\begin{array}{l} p^h > pf^h > ff^h \\ t^h > t\theta^h > \theta\theta^h \\ k^h > kx^h > xx^h \end{array}$$

I assume that the segments involved retain their |spread| specification throughout as there is no philological evidence that it is ever lost and the present-day reflexes of the HGCS₁ fricatives contrast with non-|spread| fricatives. I further assume that the two processes illustrated here are best analysed in terms of the proposals developed in section 3.2.2.3.1. This means that the acquisition of |frication| which occurs in the affrication is due to acoustic reanalysis and that the deaffrication, or spirantisation, is due to the loss of |occlusion| in the affricate. There are two main remaining issues when the HGCS₁, as formulated in (4.15), is considered as a lenition: (i) what exactly were the patterns of inhibition, and (ii) how is the geminacy of the fricatives to be accounted for? The bulk of the remainder of this section deals with these questions, but first I consider two recent alternative analyses.

In a far-reaching analysis of Germanic consonantal developments, Vennemann (1984, 1994) proposes what he calls a 'Bifurcation Theory' of Germanic consonantism. We have already encountered certain aspects of Vennemann's proposals, and several of these are compelling. His proposal for the HGCS₁ is less persuasive, however. It is, in fact, a consideration of the innovation of this process which leads Vennemann to name his proposal the Bifurcation Theory, because, as was briefly mentioned above, he proposes that the Germanic dialects split into 'High Germanic' and 'Low Germanic' dialects according to what occurred to the IE *M* series. Vennemann assumes a version of the 'Glottalic Theory', so he reconstructs the IE *M*s as ejectives. He presents the bifurcation into High Germanic and Low Germanic as in (4.16), where the top line represents High Germanic and the bottom line Low Germanic; I retain his original symbols, which can be quite easily translated into those used here.

(4.16)



Vennemann (1992, 274)

Vennemann proposes that this is a more parsimonious account of the changes, because it does away with what we have formulated above as the GCS₃ in the history of High Germanic. His definition of High Germanic and Low Germanic are as follows: “[t]he High Germanic languages include all High German dialects and the extinct Lombardic, the Low Germanic languages all remaining living Germanic languages and the extinct Gothic” (1992, 272). However, it does not seem clear to me that the account is any more parsimonious than the account which has been presented here. The GCS₃ occurs on both accounts, so it is not the case that the Bifurcation Theory truly removes the need for any diachronic process, it only removes it for ‘High Germanic’. High Germanic is essentially just High German and the closely related Lombardic, and it does not seem too problematic to assume that the GCS₃ affected the whole of Germanic as a common Germanic development, especially given the fact that Vennemann assumes that it has occurred in dialects of West Germanic, North Germanic and East Germanic, that is, in all three branches of Germanic. Vennemann’s account also rests on an assumption that it is natural for ejectives to become affricates and Vennemann (1984) proposes that this should be the case. However, the most comprehensive survey of the synchronic and diachronic phonological behaviour of ejectives, Fallon (1998), does not include a single case of the affrication of ejectives, and this might be taken to speak against Vennemann’s account.

In section 3.2.2.3.1, I proposed a straightforward account of how aspiration can lead in a natural and minimalist way to affrication, which is backed up with cross-linguistic support. It therefore seems reasonable to reject this aspect of Vennemann’s proposal (although, as we have seen above, many of the other proposals in his work on the topic are much more compelling). While it assumes one further process in the history of High German dialects, the account proposed here seems more likely.

Another recent analysis, various aspects of which are discussed in Davis & Iverson (1995) and Davis, Iverson & Salmons (1999), is closer than Vennemann’s to the analysis presented in this thesis, but differs in the means through which affrication is modelled. This analysis assumes that the Proto-Germanic inputs were aspirated, as is assumed here, but as we saw in

section 3.2.2.3.1, the authors see no way in which aspiration can easily lead to affrication. They propose a series of chronologically ordered processes, as shown in their derivations of OHG *offan* ‘open’ and *slafen* ‘sleep’, as shown in (4.17):

(4.17)

PRE-OHG	*o.p ^h an	*slāp ^h an
WEIGHT LAW	*op.han	—
SEGMENTATION	—	*slā.phan
ASSIMILATION	*op.fan	*slāpfan
WEAKENING	of.fan	slā.fan ~ slaf.fan

Davis, Iverson & Salmons (1999, 184)

The analysis here relies on the segmentation of the aspiration from the stops to become a full segment, which is done by two different means. The first of these, as can be seen in (4.17), produces a rather unusual syllabification, in which one syllable is closed in ancestors of words like *offan* by the stop and the second is opened by the newly segmented glottal fricative. The oddity of the syllable contact is partly removed on this account by the spreading of place features from the following stop onto the glottal fricative. Finally, a ‘weakening’ or lenition stage is required to remove the stop and to produce the geminate fricative, in the relevant environments (in words such as *slafen*, with long vowels, the geminate is claimed to be further simplified).

While this account works in its particulars, and achieves what it sets out to do, it seems to me that the analysis proposed in this thesis is simpler and more natural. In his general discussion of the type of ‘segmentation’ which Davis, Iverson & Salmons assume, Fallon (1998), who calls the process ‘fission’, claims that it seems to be extremely uncommon in languages, so it is not obvious that we can assume it as a natural stage in the history of German. Additionally, it is not clear that the ‘weakening’ process is motivated, given what we have established for such lenitions of stop to fricative in this thesis. The analysis proposed here requires less unnatural processes to be assumed, and, as we saw in section 3.2.2.3.1, it ties in the previous [spread] specification with the affrication, so I believe it is preferable to both the analysis proposed by Davis, Iverson and Salmons and to Vennemann’s analysis discussed above.

The two-stage analysis proposed in this thesis, is not the whole story, however. In section 3.2.3, many aspects of the inhibition of the processes were discussed, and I do not repeat here all that was discussed there. It is clear, however, on any analysis, that the initial affrication

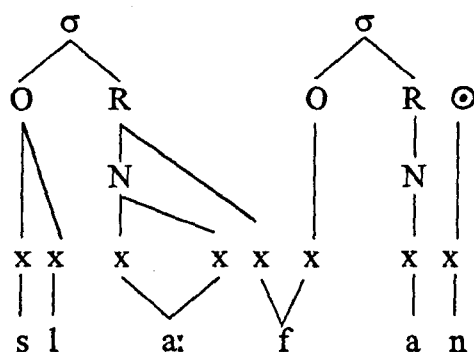
was inhibited in partial laryngeal geminates (ie, in [s__]) and, in some dialects, in partial place geminates. As mentioned in section 2.1.2, one of the cases of total inhibition was in the environment /t^hr/, and it may well be that this can be accounted for by the sharing of |coronality| between the stop and the rhotic. Davis & Iverson (1995) assume this type of account, in the form of Hayes' autosegmental Linking Constraint. We saw in section 3.2.3, however, that, although Hayes' constraint was the initial formulation of the autosegmental 'sharing' insight, it situates the explanation wrongly. I propose that in the /t^hr/ clusters, as elsewhere in melodic lenition inhibition, the sharing of elements gives 'strength' to the segments involved, which do not therefore, undergo the lenition processes. As was explained in section 3.2.3, there was notable variation in terms of melodic inhibition among the dialects. The spirantisation was inhibited more substantially. It was prosodically inhibited word-initially and by several melodic environments, where the affricates remained. In fact, the spirantisation only occurred widely in the 'intervocalic' and 'final' environments.

The spirantisation, as was noted in section 2.1.2, and also in the current section, is not typically described as a simple spirantisation, however. The account above, most recently given in (4.15) assumes that the fricative outputs of the second stage of the HGCS₁ were geminates. As was noted in section 2.1.2, this would, in fact, probably disqualify the process as being a lenition on many definitions of what lenition is (although this is not so clear on the definition adopted in this thesis). A more substantial problem is that this change, from a stop to a geminate fricative (through an affricate stage) is probably not a possible phonological process. All analyses of the HGCS₁ assume that the fricative output was a geminate because the segment is typically represented by two letters in the manuscripts, for example <offan> 'open', <la33an> 'let', <rihhi> 'country' (taken from Penzl 1975 and Davis & Iverson 1995).

One problem with the traditional analysis which sees geminates as the output of the HGCS₁ spirantisation is that in certain words, the occurrence of geminates would violate the principle of possible rhyme structure which was discussed in section 3.1.4.3, where it was named '*μμμ'. This principle states that it is not possible to have more than three timing slots in a rhyme, but, on the traditional 'geminate fricative' analysis of the HGCS₁, rhymes would be produced which would violate this principle. These would be in all those polysyllabic words (of which there were a substantial number) with a long vowel in the first syllable, followed by a single IE *T*. This includes words such as <la33an> and <slaffan> given above. In both of these words, the first, tonic vowel was long and if this is followed by a geminate, the

rhyme would contain three timing slots (or moras), as shown in the syllabic representation for <slaffan> in (4.18).

(4.18)



There is no problem for words which have a geminate after a short vowel, of course, such as <offan>, because these would only involve two timing slots in the tonic syllable, but the same problem as in <slaffan> applies to <lazzan> and all other words with long vowels.

We saw in section 3.1.4.3 that there is a substantial phonological reasoning to assume that *μμμ is a restriction on possible rhymes. Even on the type of reasoning encountered in OT, where such constraints can be violated, the violation of constraints incurs asterisks and hence would make the candidate output worse than one which does not violate *μμμ. In any case, it seems that rhymes with three timing units are not natural in phonology, and this makes the formulation of the HGCS₁ with geminate fricatives outputs even more problematic because, as we saw in the discussion of naturalness in section 1.2.2.2, the one area where there is general agreement on the role of naturalness in phonology is that it plays a role in process innovation. So we do not expect endogenous innovations to produce unnatural phonological objects. There is no evidence that the HGCS₁ was not endogenous and yet the traditional analysis proposes that unnatural objects were created by it. The problem is not just theoretical. Kirchner (1998, 2000) makes the inductive generalisation, after the extensive investigation of lenition patterns, which was described in section 3.2.2.2, that “[n]o process converts a stop (geminate or otherwise) to a geminate with reduced oral constriction” (2000, 511), so there are no attested analogues for a process such as the HGCS₁ with geminate fricative outputs.

The traditional analysis of the HGCS₁, which is shared by all recent analyses, thus seems highly problematic. Both the phonological reasoning which has led to the recognition of *μμμ and Kirchner’s empirical observations point towards the conclusion that it cannot be right. There is a solution to this problem, however. It is, in fact, the case that there is no indisputable

evidence that geminate fricatives ever occurred in the environments described above. It is generally proposed that, shortly after the innovation of the fricative stage of the HGCS₁, the fricatives shortened after a long vowel. Those present-day dialects which still have geminates show this pattern, ie, geminates can only occur after short vowels, so the traditional analysis is forced to assume a shortening. I propose that the opposite occurred. The HGCS₁ becomes an entirely natural lenition if we assume that the fricative outputs were singletons. This is what we would expect for a lenition. I propose that this was followed chronologically by a lengthening, or gemination process for those fricatives where it was possible (eg, following short vowels, but not following long vowels). Geminations of this type are well attested in the history of Germanic and other languages and so it is far from impossible that such a process could have been innovated into High German dialects after the HGCS₁.

This assumption requires a slight reinterpretation of the philological evidence, but it is not incompatible with it. As is widely noted (eg in Penzl 1971, 1975 and Davis & Iverson 1995), the segments in question were not always written with two letters in words with long vowels. In any case, the HGCS₁ was prehistoric, and we may assume that the gemination was innovated before OHG was written also (this is often assumed for the shortening of the traditional account), so it seems that the scribes were not necessarily indicating the length of the fricative by using spellings with double letters.

An important part of my account is that the fricative outputs of the HGCS₁ did not merge with the already existing fricatives of Old High German. As for example Keller (1978) shows, there was one series of fricatives in OHG. I assume that these were phonologically distinct from the newly created HGCS₁ fricatives, and there is philological support for this assumption. As, for example Penzl (1971) and Schmidt (1984) explain, the inherited fricatives are often written distinctly from the newly created HGCS₁ fricatives, thus for example, the old labial is written with the letter <u/v>, for example in OHG *ouan* 'stove' (which is spelt *Ofen* in Present-Day German), whereas the new HGCS₁ was spelt with <f>. We have already seen in section 2.3.1 that the coronal HGCS₁ output fricative did not merge with the inherited coronal fricative; the HGCS₁ output, which we have identified as /θ^h/, was consistently spelt <ʒ> or <z>, whereas the inherited fricative was spelt <s>. It thus seems that the outputs of the HGCS₁, even as singletons, did not merge with the inherited OHG fricatives and hence formed a natural

class²¹ which could undergo a gemination without affecting the inherited fricatives.

The account proposed here is more parsimonious than the traditional account because it does not assume that geminate fricatives were first created following long vowels and then lost again through shortening. The only geminates which were created are those which remained until the general loss in gemination in many varieties of German. It is therefore neither attractive, philologically necessary, nor phonologically possible to assume that the fricative outputs of the HGCS₁ were geminates and I therefore propose that the alternative account described here is the correct one. The account gives us a final version of the HGCS₁ which is given in (4.19):

(4.19) HGCS₁: final version

$$\begin{array}{l} p^h > pf^h > f^h \\ t^h > t\theta^h > \theta^h \\ k^h > kx^h > x^h \end{array}$$

In addition to the processes described in (4.19), we need to recognise a gemination, which lengthened the natural class of /f^h, θ^h, x^h/ into the second slot in rhymes, where this was possible. Additionally, along with all other analyses, we can recognise that there was an unconditioned merger between the slit alveolar fricative output of the HGCS₁ and the inherited Germanic alveolar fricative. Russ (1982) dates this to around the 13th century.

These changes combine with those which I have shown in this section to be compatible with both the insights of phonological theory and with the philological record, to provide a full picture of the segments involved throughout the history of German. I turn now to the HGCS₂, which, as we will see, is rather different.

4.1.2.2 HGCS₂

The HGCS₂ seems to involve a 'devoicing' in traditional terminology. This might well be problematic for the position developed here as it would likely have to involve the unmotivated acquisition of |spread|. The process as formulated in (4.20) shows that the inputs were neutral stops, the outputs of GCS₂. We saw in section 3.1.3.1 that such segments can be written as either <b, d, g> or <p, t, k>, depending on whether the language is a |spread| or |voice| language.

²¹ It seems most likely that the contrast would be based on the segments' laryngeal elements. As the HGCS₁ fricatives were |spread| fricatives, we might assume that the inherited fricatives were neutral fricatives. This would entail their having lost their own |spread| specification previously, as some of these inherited fricatives were the outputs of the GCS₁.

It seems that the only way in which scribes would start to write such segments as ⟨p, t, k⟩ when previously they had been written ⟨b, d, g⟩ would be if they acquired [spread], which would mean that the HGCS₂ would be as in (4.20):

(4.20) HGCS₂: take 3

$p^o > p^h$

$t^o > t^h$

$k^o > k^h$

However, as we will see in this section, at least some of the details of these processes can be seen to rather melt away when they are considered in detail and there may well be another way of interpreting the change in spelling.

One clear point to note is that the philological evidence for the HGCS₂ is much less certain than it is for the other processes that are discussed in this thesis. As was noted in section 2.1.2, the situation in the OHG manuscripts is often that there was variation between the two sets of letters, rather than it being the case that the letters ⟨p, t, k⟩ were always used. For example, Keller (1978) described a “widespread fluctuation in spelling ... especially in Upper German documents” (1978, 173) particularly between ⟨p⟩ and ⟨b⟩ and between ⟨k⟩ and ⟨g⟩.²² In fact, the process is often described (for example in Paul 1944) as having only affected the coronal stop, because there is not even substantial evidence of variation for the labial and velar in manuscripts from most areas (apart from the Upper German examples of the type that were given as data in section 2.1.2). This may be taken to mean that no such ‘devoicing’ process occurred, at least for the labial and velar and at least in most varieties.

One further point to bear in mind is that if a language (or dialect, if a distinction is made) has only one series of stops (or has no contrast at one place of articulation), then there is no need to make a distinction in spelling between ⟨b, d, g⟩ and ⟨p, t, k⟩. Convention may dictate that one set should be used, but, especially at stages in the history of a language where there is no clear standard for spelling, it does not matter whether the one stop at any place of articulation in a phonological system is written with a letter from either series of letters. Thus /p^o/, for example, could be written ⟨b⟩ or ⟨p⟩. It could be that it is this type of situation which led to the variation found among the letters used in the manuscripts to spell the stops of OHG and MHG. It is notable that the Upper German areas, where most of the variation in spelling

²² I do not consider the geminates here as they are not strictly relevant to the spelling of the singletons.

occurs, are those most affected by the HGCS₁, which removed many of the stops in the variety, as we saw in section 4.1.2.1, by turning them into affricates and fricatives. If only one series of stops is left (either totally or in certain phonological environments) then the variation in spelling which is found in certain manuscripts can be interpreted as just that: variation in spelling, which does not indicate a phonological change.

We will see in section 4.2 that the German reference variety *Standardlautung* has not developed straightforwardly from High German varieties, so we cannot necessarily project back from Present-Day standard pronunciations to past phonological states. We will further see in section 4.1.4 that the IGCW may be relevant here, too, in helping to explain the variety in spelling that exists among the stops in the older German manuscripts.

The discussion in this section has been quite brief, but it is intended to show that the status of the HGCS₂ is rather uncertain. While it may be that we need to recognise a process of the type give in (4.20) for certain segments in certain varieties, the true picture may well simply be that an absence of laryngeal distinctions (ie, the existence of only one series of stops) led to the observed variation in spelling. The uncertainty leads to a final version for the HGCS₂ which illustrates the hesitancy with which I believe we should approach the process. The development is certainly less secure than the HGCS₁, and this recognition is partly due to the reasoning that led to the recognition of laryngeal realism. We saw in section 3.1.3.1 that great care must be taken in interpreting the spelling of stops.

(4.21) HGCS₂: final version?

no change?

In the next sections, we return to more certain phonological innovations. We will see that some of the reasoning developed in this section is relevant to aspects of the understanding of the EIFV and the IGCW. The two processes will, in fact, be seen to be rather similar.

4.1.3 The English Initial Fricative Voicing revisited

Compared to the GCS and the HGCS, the EIFV has received much less analytical attention in the literature, something which it has in common with the IGCW (which, as we will see in the next section, is unfortunate given the potential importance of the data discussed there). Much of the attention that has been directed towards the EIFV has been to do with the philological points which were discussed in section 2.1.3, when the process was first introduced. These focus not so much on what occurred but on when and where it occurred.

We saw in section 2.1.3 that the effects of the process were once widespread throughout the south of England and that one tradition of dating proposes that the process was first innovated well after the Anglo-Saxon settlement of Britain (Brunner 1965 and Fisiak 1984 tend towards this opinion, for example). The other tradition of dating proposes that the process was a shared West Germanic innovation, because there were somewhat analogous processes in other Germanic languages (Bennet 1955 and Lass 1991-1993 defend this position, for example). As we further noted in section 2.1.3, the precise dating and situation of the innovation of the process is not of paramount concern here, although I do make some proposals on this issue below. Wherever it was first innovated, the input to the process was largely the single Germanic series of fricatives. Most of these were the outputs of the GCS₁, although the alveolar was inherited from IE. In Proto-Germanic, and hence in all Germanic languages, at least initially, the fricatives were |spread|, as we saw for the outputs to the GCS₁ in section 4.1.4. The dialect data given in the initial discussion of the process shows that it also affects the novel segment /ʃ/, which was derived in English from the Germanic /sk/ cluster. This is given in the version of the process in (4.22), which is largely the same as ‘take 2’ from (2.33), except that the inputs have been altered to fit in with laryngeal realist assumptions, as was the case in the initial discussion of the GCS and HGCS above.²³

(4.22) The EIFV: **take 3**

f^h > v

θ^h > ð

s^h > z

ʃ^h > ʒ

The process is described in its title as being ‘initial’, but this is misleading, as we saw in section 2.1.3, because the fricative segments in all OE dialects were also subject to a process of ‘medial voicing’ so that the process described in (4.22) was quite possibly largely uninhibited. It will certainly be clear from the process’s name, as well as from the data given in section 2.1.3 that there was no prosodic inhibition in the lenition. Lass (1991-1993) reports that there was no lenition in words such as *speche* ‘speech’, so it is possible that partial laryngeal geminates inhibited the lenition.

²³ For the inherited alveolar, I assume that it either already had |spread|, in IE, or acquired it in conformity with the other fricatives of Germanic which had been derived by the GCS₁. It seems quite clear that /ʃ/ would also be a |spread| segment as the two segments which it derives from were both |spread|. Initial Germanic /x/ had debuccalised to /h/ before the process was innovated (see sections and 2.1.1.5 and 3.1.3.1.8), so was not a possible input to the EIFC.

The formulation of the process in (4.22) is not the only possible interpretation of the process, however, and I propose that there is a more natural and minimalist analysis which will allow us to unify the process here with a range of other processes which are common types of lenitions. On laryngeal realist assumptions, the version in (4.22) requires us to assume that the fricatives pick up |voice| as well as losing |spread|, but this is not a necessary assumption, in fact, there are good reasons to believe that this was not the case. A simpler process would involve the fricatives simply losing their |spread| element, in a case of delaryngealisation, as discussed in section 3.2.2.3.4. The result of delaryngealisation is neutral segments, which, as we saw in the last section and elsewhere, can be written with the letters which are used to write truly |voice| segments in other languages. Perhaps more importantly here (since much of the evidence for the process is in the form of orally attested pronunciations from non-standard dialects), such segments are clearly distinct from |spread| fricatives and the distinction between |spread| and neutral segments is proposed above to be the basis of the contrast in stops. It is therefore no surprise that the neutral fricatives of the southern varieties, where |spread| has been lost, should be perceived as the type of segment which should be written with such letters as <v, z> and not with <f, s>. It is also clear that present-day speakers of non-EIFV varieties will identify the fricatives of the present-day EIFV varieties with their own neutral fricatives, which now contrast with |spread| fricatives. The analysis of the EIFV as delaryngealisation is given in (4.23):

(4.23) EIFV: final version

$$\begin{aligned} f^h &> f^o \\ \theta^h &> \theta^o \\ s^h &> s^o \\ \int^h &> \int^o \end{aligned}$$

This analysis has the advantage that it only requires one quantum, unlike that in (4.22) and that it goes some way to explain why changes of this type seem so common in Germanic languages. We can assume that the ‘standard’ OE medial ‘voicing’ was also a case of delaryngealisation, and that the EIFV has, as a consequence, that this process of medial delaryngealisation is lost from the phonology of the EIFV varieties because, once the EIFV is lexicalised, it is clear that the underlying segments will be neutral fricatives.

We saw in section 3.1.3.1.11 that the laryngeal realist position, which was originally developed to account for stops, can also be applied to fricatives and we saw that recent work by Vaux (1998) especially, has shown that what have previously simply been referred to as

‘voiceless’ fricatives can, in fact, be characterised by [spread]. This opens up the possibilities of analyses such as the one proposed here for the EIFV. On this picture, the change seems very natural, as it involves the loss of an articulatory gesture, and we might speculate that such a process might be particularly likely where there is no need to retain the laryngeal element in order to maintain a contrast, as is the case here for the EIFV, because there was only one set of fricatives before the process, and still only one afterwards.

On the reasoning developed here, the innovation of the process is no real surprise. This means that we should not wonder if the process is innovated in languages with a similar phonology, and it may be that this realisation makes the proposal that the EIFV is really a common Germanic innovation less likely. It certainly makes it seem less necessary to assume that the similar processes in other West Germanic dialects (described, for example in Nielsen 1981) are the result of a shared innovation, as the scenario of parallel innovation seems entirely possible, as discussed in section 1.3. We will see below that the type of process envisaged here is indeed quite common.

4.1.4 The Inner-German Consonant Weakening revisited

The last of the four sets of Germanic data has perhaps been the least discussed up till now in this thesis. This may be slightly surprising, given the fact that it has the potential to upset one of the most widely accepted ‘phonological universals’, as we saw in its initial presentation in section 2.1.4. This lack of attention is, however, commensurate with the way in which the process has generally been neglected in phonology. The IGCW was described in section 2.1.4 as a type of ‘voicing’, which affected the stops in a number of High German dialects. These stops were those *T*s which had been inherited from Proto-Germanic, as outputs from the GCS₃, but had not been affected by the HGCS₁ (ie, those in environments where the HGCS₁ processes were inhibited), and also any *T*s which occurred in words that had been borrowed since the HGCS₁ (from other languages or from German dialects which had not taken part in the HGCS₁). As we saw in the discussion of the inhibition of the HGCS₁ processes in sections 2.1.2.3, 3.2.3 and 4.2.1, in certain non-southern dialects, the HGCS₁ processes were entirely inhibited in several environments, especially in the segments /p^h/ and /t^h/. I give an interpretation for the IGCW in (4.24). This is reformulated from (2.22) to take account of laryngeal clarity for the inputs, but not the outputs, as has been conventional in this chapter.

(4.24) IGCW: take 2

$p^h > b$

$t^h > d$

$k^h > g$

As was explained in section 2.1.4, the process was inhibited differently in the dialects into which it was innovated, and in certain varieties, it seems to have been inhibited prosodically in word-initial position. In other varieties, chief among which are the South Osterländisch varieties of Upper Saxon (see Albrecht 1881, Bergmann 1991), the lenition was entirely uninhibited. The import of this process for wider phonological theory derives from the fact that, as is relatively clear from the formulation in (4.24) and as was made explicit in (2.23), the IGCW actually represents an unconditioned merger in the stop system of the German dialects into which it was innovated. In the South Osterländisch varieties, this means that there was only one series of stops left in the phonological system. The problem alluded to at the start of this section is due to the interpretation of the output stops given in (4.24). As shown there, they stand for fully voiced stops, which are characterised by the laryngeal element [voice].

It is far from phonologically 'impossible' for languages to have only one series of stops. Maddieson (1984) reports 50 cases, which is 15.8% of his corpus, but these all have voiceless plosives.²⁴ If we return to the quotation from Hyman (1975) in section 1.2.2, it seems that the IGCW, as given in (4.24) is a counterexample to a well established phonological universal, which is backed up by Maddieson's survey. The relevant part of Hyman's quotation is given below:

...a sound change turning all instances of [p, t, k] into [b, d, g] has never been reported. If such a sound change were to take place, the resulting system would include a series of voiced stops but no series of voiceless stops. In other words, the Jakobsonian implicational universal whereby /b, d, g/ implies /p, t, k/ would be violated. As pointed out by Greenberg [(1966, 510)], any sound change which produces an impossible sound system (such as the one which would result from a change voicing all voiceless stops) is an impossible change.

(Hyman 1975, 17-18)

There is no reason to believe that the IGCW was not endogenously innovated and so, it seems that either the phonological universal is wrong, or the IGCW did not exist. We might entertain the latter idea, given the discussion in section 4.1.2.2, but this is not a possibility

²⁴ Maddieson explains that there is good reason to believe that the one language which has been claimed to have a single series of voiced stops, Bandjalang, has been incorrectly reported.

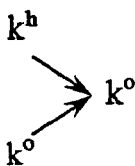
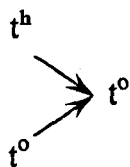
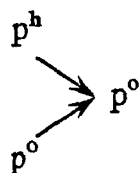
because there are present-day dialects which show the effects of the unconditioned merger, as reported, for example, in Schübel (1955), Bock (1965), Bergmann (1991) and Weldner (1991). In fact, it may be that the effects of the IGCW have led to the possibility, which was entertained in section 4.1.2.2, that the HGCS₂ has been misinterpreted. Once there is only one series of stops in a system then it does not matter whether they are spelt with the series ⟨p, t, k⟩ or ⟨b, d, g⟩. This is illustrated by the fact that there are many reports of the confusion of these two orthographic series by speakers of dialects which have been affected by the IGCW (as explained in section 2.1.4, and see Ammon & Loewer 1977, Zehetner 1977 and Kraemer 1978).

The IGCW clearly did occur, and this puts Hyman's (and Jakobson's) phonological universal under threat. The process seems to be an example of an impossible phonological innovation. There is, however, another possible conclusion, thanks to the laryngeal realist interpretation of laryngeal phonology which was developed in section 3.1.3.1. This relies on the recognition that the inputs to the IGCW were |spread| stops, and on the identification of delaryngealisation as a common type of lenition process (as discussed in section 3.2.2.3.4). With these assumptions, we can analyse the IGCW as a simple case of the loss of |spread|. This means that the outputs of the process are neutral stops, not |voice| stops, and the Hyman's universal is rescued, but only if the segments which he refers to, using *M* symbols, are interpreted as |voice| stops, and not neutral stops (which they are often used to transcribe). The universal can then be recast to recognise that an unconditioned merger involving $p^h, t^h, k^h > b, d, g$ is impossible, as is $p^o, t^o, k^o > b, d, g$, but $p^h, t^h, k^h > p^o, t^o, k^o$ is fine, and we might predict that $b, d, g > p^o, t^o, k^o$ can also occur. This might be seen to make the 'neutral' stop the unmarked case, as has indeed been proposed, and it then becomes quite expected that languages should feature the neutral series and this, in turn lends further credence to the laryngeal realist analysis of IE languages into |spread| languages and |voice| languages, because both types would feature the unmarked series. On the 'traditional' non-laryngeal-realist view, none of the IE languages make use of the unmarked series.

The analysis proposed here for the IGCW is only possible if we recognise the three types of phonological object which are used in the laryngeal realism (|spread|, |voice| and neutral stops). This has been recognised before (indeed the outputs of the IGCW are often described as 'lenis' rather than 'voiced', in line with the discussion in section 3.1.3.1.4), but has not been integrated with phonological theory or with the general universalist concerns discussed here.

The final version for the IGCW therefore recognises that the outputs were neutral stops, as they are in many non-reference varieties of German today. We may also note that this makes the segmental merger with the other series of stops quite expected because they were neutral stops, too. These were the output of the GCS₂, which, as we saw in section 4.1.2.2, were probably not affected by an HGCS₂ process. The final version of the IGCW is given in (4.25). It is shown to illustrate the merger which occurred, although it will be clear that the inherited neutral stops were not affected themselves by the process.

(4.25) IGCW: final version



This analysis means that we do not have to assume that the *Ms* were altered at all, which makes the analysis more parsimonious than that of Schieb (1970), for example, who proposes that “in many areas of Upper and Central German, the fortis segments *t*, *p*, *k* were lenited to voiceless lenis *d*, *b*, *g*, which merge with the formerly voiced stops which lose their voicing.” (1970, 367).²⁵ There is no need to assume an additional process which causes the *Ms* to lose [voice], because, on the laryngeal realist analysis developed here, the Germanic *Ms* were neutral stops.

On this analysis the IGCW is shown to be essentially the same kind of delaryngealisation process as the EIFV, in that both involve the loss of [spread]. This process is also very similar to the common Germanic part of the GCS₂, where [spread] was lost along with [voice],

²⁵ The original reads: “In wieten Teilen des Obd. und Md. erleiden die Fortes *t*, *p*, *k* eine Schwächung zu stimmlosen Lenes *d*, *b*, *g*, die weithin mit den ehemals stimmhaften, die umgekehrt ihren Stimmtön verlieren, zusammenfallen.”

although neither the EIFV nor the GCS₂ involved a loss of contrast. The IGCW, on the contrary, involved a complete collapse of laryngeal distinctions, which had a substantial effect on the underlying phonological system. One thing that all three of these processes have in common is that they all involved the creation of laryngeally neutral segments and this is a vital piece of our understanding of them - it is this which allows us to interpret them as possible phonological innovations.

In his treatment of the EIFV, Lass (1991-1993) discusses the kind of phonological dilemma that we have dealt with in detail in this section. He sees the EIFV as problematic because he assumes that the fricatives are characterised by [voice]. He writes that the EIFV produces:

...a fricative system with only one lexical or non-derived glottal state, but this time voiced rather than voiceless. Is this legal? The answer seems to be no. According to the best recent survey (Maddieson 1984: ch2), there don't appear to be any languages like this; and on the general uniformitarian principle that we do not reconstruct for the past any *état de langue* that is in principle impossible at present...

Lass (1991-1993, 29)

The analysis which was presented in section 4.1.3 shows that this is not, in fact, a problem, because the segments are interpreted as laryngeally neutral, not as [voice] segments, just as was the case for the IGCW.

This wider section (4.1) has involved the application to concrete data of the theoretical analyses which have been developed throughout this thesis, such as the elemental approach to subsegmental structure, combined with laryngeal realism, and the understanding of prosody, of naturalness, minimality and notions of what is possible in phonology in general and lenition in particular. We have seen that these ideas allow for often novel analyses of the four sets of Germanic data which were first introduced in chapter 2. I believe that the analyses proposed here allow a deeper insight into the processes involved than has previously been possible. Some further general conclusions about the analyses that have been proposed in this chapter are drawn in chapter 5, along with a wider final consideration of the issues which have been addressed in this thesis. Before we embark on this conclusion, however, section 4.2 deals, quite briefly, with certain other important aspects of the phonology of (West) Germanic languages which have become live issues thanks to the theoretical positions developed in this chapter and elsewhere in this thesis.

4.2 The laryngeal history of Present-Day German and Dutch

In this second part to chapter 4, which is much shorter than the first, I discuss aspects of the laryngeal phonology of two present-day Germanic languages. The analyses proposed here deal with aspects of the phonology of Present-Day German and Present-Day Dutch, which have already been mentioned at various points in this thesis. The treatment here is intended to round off the story as far as they are concerned. It will be shown that the present-day situation in both languages is entirely compatible with the points which have been developed in earlier chapters. I deal with the diachronic origins of the laryngeal distinction in the Present-Day *Standardlautung* variety of German in section 4.2.1 and I deal with the nature and diachronic origin of the laryngeal distinction in Present-Day Dutch in section 4.2.2.

4.2.1 The problem of standards: the case of German

In section 3.1.3.1, the phonetics of *Standardlautung* were considered in some detail, largely thanks to the work of Jessen (1997, 1999). It was shown there that the phonetic properties of the stops of *Standardlautung* are not compatible with the ‘traditional’ phonological analysis of them, which describes them as featuring [voice] (in fact the traditional distinction is between [+voice] and [–voice], as we saw, using a binary feature). Jessen shows convincingly that a different type of analysis is required, which is more compatible with the philological distinction fortis/lenis. In section 3.1.3.1, I formalised this as ‘laryngeal realism’, recognising [spread] languages (such as *Standardlautung*, most varieties of English, Icelandic, Danish and the like) and [voice] languages such as (Spanish, Russian, French and Dutch). This was the basis for much of the discussion in section 4.1. As part of section 4.1, however, we saw that many of the originally [spread] stops of Germanic became affricates and fricatives in varieties of German (in the HGCS₁), and that the element [spread] was lost (in the IGCW) in many of the remaining stops in remaining varieties of German. The question thus arises as to where the [spread] stops that Jessen and others have analysed came from. In this section, I sketch an outline of the history of *Standardlautung* which will allow us to answer this question. The account will illustrate the points made in section 1.2.3 about the relationship between ‘standard’ or ‘reference’ varieties and ‘traditional dialect’ varieties of a language. The history is a quite standard one, but it has been recently investigated in detail by Schmidt & Vennemann (1985a,b), who find that there is good reason to accept it.

As Schmidt & Vennemann (1985a,b) note, along with many others, for example Keller (1978) and Russ (1982), a quite standard assumption in German philology is that

Standardlautung is not easily tied to any traditional dialect of German. This is not uncommon in the history of standard forms of languages, and such varieties may not always develop naturally (in the technical sense). They are often affected much more by exogenous factors than are traditional dialects (such as Upper Saxon and High Alemannic in Germany and Northumbrian and Yorkshire in Britain). If any of the dialects which were affected by the IGCW had formed the basis of *Standardlautung* by themselves, then the reference variety of German would have only one series of stops, at least at certain places of articulation, and these would be neutral stops. This is clearly not the case.

Traditional dialects are still spoken in Germany, probably more so than in Britain, but *Standardlautung* is now spoken in many areas of the country, at least by a certain section of the population (see, for example, Barbour & Stevenson 1990 for some discussion of these points). *Standardlautung* is not a natural further development of the traditional dialects of these areas, but is a phonology which is being exogenously adopted through sociolinguistic pressures. The origin of *Standardlautung* is generally believed to lie in the imposition of the phonology of northern, Low German onto the syntax and morphology of southern High German. As Schmidt & Vennemann (1985a,b) note, this insight may first have been expressed by Viëtor (1893), and it relies on the recognition that, even in those varieties of German where the IGCW was innovated, a distinction was often maintained between two series of stops in spelling. We have seen in section 4.1 that this is not so clear at earlier stages of German, but in the later period, there was a relatively well established standard of spelling

As Schmidt & Vennemann (1985a) explain:

While the central dialects lost the distinction between the fortis and lenis series ($p - t - k$ vs. $b - d - g$) ... the historical orthography of the written language kept this distinction in the corresponding series of letters $\langle p\ t\ k \rangle$ vs. $\langle b\ d\ g \rangle$...

Schmidt & Vennemann (1985a, 165)²⁶

The introduction of $[\text{spread}]$ and the general transformation of *Standardlautung* into a $[\text{spread}]$ language came through contact with Low German. As Schmidt & Vennemann (1985a,b) show, basing their claims on such work as Brugge (1944) and Peters (1973), Low German is clearly a $[\text{spread}]$ language, with distinct aspiration in the *Ts*. Low German has not undergone such processes as the HGCS or the IGCW and so, like English, Danish and

²⁶ The original reads: "Während die zentrale Dialekte den Unterschied zwischen der Fortis- und der Lenis Reihe ($p - t - k$ vs. $b - d - g$) ... aufhoben, bewahrte die historische Graphie der geschriebenen Sprache diesen Unterschied in den korrespondierenden Graphenreihen $\langle p\ t\ k \rangle$ vs. $\langle b\ d\ g \rangle$ "

Icelandic, it maintains the original Germanic opposition of |spread| *T*s contrasting with neutral *Ms*. Low German speakers started to speak High German as a second language, due to the unification of Germany as an empire and the political promotion of High German at the expense of Low German. When they spoke High German, they naturally transferred the contrast which they made between series of stops in Low German to the contrast which was spelt in the High German and therefore pronounced those High German segments which were spelt <p, t, k> as /p^h, t^h, k^h/ and those spelt <b, d, g> as /p^o, t^o, k^o/, as was the case in Low German writing.

This form of pronunciation became widespread throughout Northern Germany, where Low German was spoken, and subsequently became the prestige variety which is now called *Standardlautung*. As Russ (1982) explains:

What we have here is not so much a phonetic change but rather a change of linguistic model, whereby the prestige pronunciation of one area has supplanted that of another area. ... [W]ith the decline in prestige of Saxony and the rise of Prussia and furthermore the rise of Berlin as the capital of the united German Empire after 1870, the North German model of pronunciation came to be regarded as the prestige model.

Russ (1982, 26)

It is this *Standardlautung*, which is based partly on spelling pronunciation, which has now spread back into the south of Germany with its |spread| specification in stops. This means that |spread| stops were lost through natural endogenous phonological innovation throughout much of Germany in the traditional dialects, but have since returned to be spoken in those areas where they was lost. In many areas of present-day Germany, therefore a |spread| language, *Standardlautung*, is spoken alongside the traditional dialects which have no laryngeal contrast in stops; often, doubtless, speakers are bilingual in these two varieties. Both systems are natural languages and the laryngeal specifications of *Standardlautung* are, in fact, inherited from Proto-Germanic and are therefore perfectly illustrative of a 'Germanic |spread| language', but this inheritance of a laryngeal specifications did not occur directly through High German descent, but indirectly, via Low German.

4.2.2 Laryngeal switch: the case of Dutch

In this final section of chapter 4, I briefly discuss the laryngeal phonology of Dutch. What is said here also applies with some adjustment to the other members of the 'Netherlandic group'

of languages, that is, Frisian and Afrikaans. Throughout this thesis, when the Germanic languages have been mentioned in connection with the considerations of laryngeal phonology, Dutch (and the allied languages just mentioned) have always been the exception. In this section I explain why this is so and address the diachronic development of the situation in Dutch, in particular.²⁷

We saw in section 3.1.3.1 that the reference variety of Present-Day Dutch (*Algemeen Beschaafd Nederlands* = ‘ABN’), along with many non-reference varieties, shows all the signs of being a [voice] language. As is widely recognised in the literature (eg, Cohen, Ebeling, Fokkema, van Holk 1972, Iverson & Salmons 1995, Lass 1997), there is no aspiration in the Dutch *T*s, and the *M*s are typically fully voiced. There is no process of sonorant devoicing in Dutch, which is frequently found in [spread] languages, and derives ‘voiceless’ sonorants when next to a *T* through the spreading of [spread] (as in English *plan* [p_hlan], *treat* [t_hri:t]). Because Dutch has neither this type of ‘assimilation to voicelessness’ nor aspiration, there is no evidence for [spread] in Dutch *T*s.

It has been claimed that Dutch features a different type of assimilation to voicelessness than that mentioned here. Booij (1995) writes, for example, that “a fricative is devoiced after a voiceless obstruent” (1995, 58). It is noticeable, however, that the fricatives involved here are syllable initial (as in, for example, *opvallend* [ɔpfulənt] ‘remarkable’, *slaapzak* [sla:psak] ‘sleeping bag’ and *dakgoot* [dakxo:t] ‘gutter’). Rather than assimilation to [spread], this can be analysed as a case of the loss of [voice]. This is a common process in many varieties of Dutch, especially in initial environments, where the segments which are typically described as underlying /v/ and /ɣ/ normally surface as [f] and [x], for example in *vuilnis* [føylnis] ‘rubbish’ and *gezag* [xəzax] ‘authority’. Lombardi (1991) shows how Booij’s ‘regressive assimilation’ can be analysed as a special case of the loss of [voice] in the fricatives concerned and so this, too, is not evidence for the presence of [spread] in the phonology of Dutch.

There is evidence for the spread of [voice] in Dutch, however, and this is not found in non-Netherlandic Germanic languages. We saw in section 3.1.3.1.5, that Dutch features a process of ‘assimilation to voicing’ which affects underlying *T*s so that they surface as *M*s when an *M*

²⁷ When I refer to ‘Dutch’ in this section, this is intended as a cover-term to include all the varieties used in the historically contiguous Dutch and Flemish speaking areas. These varieties are now, of course, spoken in present-day Belgium and the Netherlands, and in other parts of the world.

follows directly afterwards (some of Booij’s 1995 examples are *klapbant* [klabbant] ‘flat tyre’, *potdicht* [paddixt] ‘tight’, *kookboek* [kogbuk] ‘cookbook’). There is no reason why this should be analysed as the loss of |spread| in the *T*s, because there are no cases of ‘spontaneous voicing’ of these segments (while there *is* ‘spontaneous devoicing’ in the fricatives just discussed). This phonological behaviour can therefore be taken as evidence for the presence of |voice| in the Dutch *M*s.

There are therefore both phonetic and phonological reasons to assume that Dutch is a |voice| language, unlike other Germanic languages (including *Standardlautung*). There is no problem in recognising this is the synchronic language state. There is no synchronic reason why Dutch should necessarily have the same laryngeal phonology as its Germanic neighbours. However, the issue becomes interesting when the synchronic situation is considered in connection with the concerns of diachrony. The study of diachrony has played an important role in this thesis and it is natural to wonder what is the diachronic ‘explanation’ for the difference in laryngeal phonology between Dutch and its northerly neighbour, Low German, for example.

We have noted at several points in this thesis (most recently in section 4.1) that Proto-Germanic was indisputably a |spread| language. The obvious implication is that the laryngeal phonology of Dutch has changed since Proto-Germanic, in the way shown in (4.26). However, as may be clear from (4.26), this is a rather unusual kind of process:

(4.26)

Proto-Germanic	>	Dutch
spread	>	voice
p ^h : p ^o	>	p ^o : b
t ^h : t ^o	>	t ^o : d
k ^h : k ^o	>	k ^o : g

The changes involved here did not simply involve one segment, nor one series of segments. All of the stops are affected by the change, and, in fact, the most insightful formulation of the process involved is that the phonological basis of the contrast has changed, rather than simply the elemental make-up of the segments. This seems to be a case of a complete laryngeal switch, rather than a simple phonological process. The language has stopped making use of |spread| in its phonology and started making use of |voice|. While it is not difficult to understand what has occurred here, it is not so clear *how* this change occurred.

We could argue for a 'chain shift'-type analysis, which might involve firstly $p^o, t^o, k^o > b, d, g$ and then $p^h, t^h, k^h > p^o, t^o, k^o$, but it was recognised above that chain shifts are not by themselves necessarily insightful analyses, because they do not give an explanation for why any of the individual processes were innovated. For example, the chain shift just proposed involves, as an initial stage, the unmotivated acquisition of [voice] by neutral stops. There seems to be no natural way of explaining this process, however, and it is rather dubious. There is also no reason why the second change should have to follow the first. In short, while the 'chain shift' analysis of the processes involved may reflect the truth of the situation, it is problematic, and I propose that there is another, simpler explanation.

The account given here is unlike much of the other discussion of diachronic phonological innovation in this thesis and this is connected with the fact that the development in Dutch, as described in (4.26), seems rather different to the types of phonological processes that have been discussed up till now (apart, perhaps, from that mentioned in the directly preceding section 4.2.1). In section 1.3, I discussed the distinction between endogenous and exogenous motivations for linguistic innovation. The lenition processes that have been discussed in this thesis are classic examples of endogenous innovations, because they can all be seen to be attributable to phonetic and phonological factors which are inherent in the phonological systems in which they occur. The simplest explanation for the case of laryngeal switching that has occurred in the history of Dutch (and Frisian and Afrikaans), however, is that it is due to exogeny.

The area of Europe where Dutch is now spoken has not changed massively in the past two millennia, since the area was first settled (see, for example, König 1978, Vekeman & Ecke 1992, Ramat 1998b) and throughout much of this time it has been in contact with Romance languages. We saw in section 3.1.3.1 that Romance is reconstructable as a [voice] language, and that Present-Day Romance languages, such as French, are [voice] languages. It may not be possible to identify the exact period in which the change occurred in the history of Dutch,²⁸ but it seems reasonable to assume that such an extensive period of contact between Dutch and Romance languages may have led to the borrowing into Dutch of the way of making a laryngeal contrast on the basis of [voice], to replace the inherited Germanic contrast, which was based on [spread]. Vekeman & Ecke (1992) show that there have been many types of contact, especially with French, which could have led to this exogenous innovation. In several

²⁸ This is because, as we have seen, the distinction between [spread] language and [voice] languages is not shown in spelling.

periods in the history of Dutch, French has been a prestige language for the Dutch-speaking communities, and the present-day linguistic situation in Belgium illustrates the close proximity of the two languages. It is not just the case that French has had prestige status in Dutch-speaking communities at several periods (for example, in Middle Dutch, the 17th and 18th centuries, as Vekeman & Ecke 1992 explain, among others), the two speech communities have long shared a common border where bilingualism would unavoidably be common.

Such language contact situations can easily lead to the areal dispersal of linguistic features, as was discussed in section 1.3, and analogous types of borrowings to the one proposed here for Dutch are well attested in other languages. For example, the spread of clicks though non-related languages in southern Africa, which was mentioned in section 1.3, involves a broadly similar type of change, in that the distinction between /ɔ, ɪ/ and /p, t/, for example, does not rely on the articulation of the front portion of the tongue, but rather on airstream mechanisms. We have seen that, similarly, the distinction between /p^h, t^h/ and /p^o, t^o/ does not rely on the articulation of the front part of the tongue, but rather on laryngeal activity. If types of airstream mechanism can be borrowed, it seems reasonable to conclude that types of laryngeal activity can be borrowed, too. We have also briefly discussed, in section 4.1, the fact that certain dialects of Armenian have undergone an entirely parallel type of innovation to that proposed here for Dutch. These dialects borrowed the way of making a laryngeal contrast on the basis of |constricted| from neighbouring Caucasian languages and now have a series of ejectives (see Gamkrelidze & Ivanov 1984).

Given the points made above, it seems reasonable to propose that the switching of the basis of laryngeal contrast in Dutch is an exogenously motivated innovation, and should not be accounted for by the postulation of phonological processes which were ever synchronically active in the language. The laryngeal contrast in Dutch was borrowed through contact with Romance languages, probably with French. We can assume further that the situation in Frisian is due to contact with Dutch. There has long been bilingualism in Dutch among Frisian speakers in the Netherlands, so it is not surprising that this aspect of phonology should be borrowed. The situation in Afrikaans is easily explained, as the language is largely diachronically derived from Dutch.

The innovation described here accounts for why Dutch and the other 'Netherlandic' languages are no longer |spread| languages. We saw in section 4.1.1 that certain varieties of German are no longer |spread| languages, although *Standardlautung* is. The reference form of

Dutch, ABN, is a [voice] language, because the traditional dialect varieties on which it is based underwent the laryngeal switch described above (see, for example, van den Toorn 1973 for an account of the formation of ABN).

This discussion of Dutch concludes the current chapter. We have seen here that the fact that Germanic was a [spread] language has been a quite important motivating factor for several of the processes which were described in section 4.1. In the current section (4.2) we have discussed two cases where exogeny has complicated the picture in different ways. In the case of *Standardlautung*, [spread] was reintroduced into one variety of a Germanic language. In the case of Dutch, [spread] was lost.

5 Conclusions and mutual implications

This thesis has tried to do several things. The main aim has been to combine aspects of phonological theory with historical data and reasoning, and to illustrate some of the results that can be achieved when this is done. The thesis has naturally had a narrow empirical focus and several issues have been left open. Indeed, it is to be hoped that some of the points made here will provide the basis for fruitful further research. I believe that some of the analyses carried out here and some of the predictions that were made open up a new way of viewing certain data and of considering what we might expect to find in it. In this final chapter, I summarise the main claims that were made in the thesis and draw out the implications of them. I show how some of the general empirical hypotheses that were formulated in the thesis can be tested and I consider very briefly some of the types of evidence that could disprove them.

Several proposals have been advanced in this thesis. Some of these are the specific analyses that were proposed for the four sets of Germanic data in chapter 4. Much of the analysis in that chapter rests on proposals that were put forward to account for subsegmental phonological structure (and certain other aspects of phonology) in section 3.1, and on proposals that were developed to explain obstruent lenition in section 3.2.

While many of these proposals are connected, certain aspects of them at least are independent of each other, so that, if they are shown through subsequent research to be at fault, some of them can fall, while others can remain. For example, if one of the analyses in chapter 4 is shown to be inadequate, this would not necessarily disprove the others, nor would it call the whole analysis of laryngeal realism into question. If an analysis of one of the types of lenition processes is shown to be faulty, then others can still be correct. I believe that the analyses proposed here are coherent, however, and that they conspire together to present a total picture of certain small aspects of phonology and of the historical phonology of certain Germanic languages. To the extent that the picture that has been painted here is coherent, then I believe that the individual parts of the analyses support each other.

If nothing else, then I hope that the discussion in this thesis has shown that, in order to understand the type of historical data discussed in chapter 2, we need to consider both philological and phonological insights. And I hope further that it has shown that such historical data, and the analysis that we make of it, can play a vital role in shaping our general understanding of phonology.

The notion of lenition has played an important role in the history of phonological discussion at times, and it has played an important role in this thesis. The notion has long had a rather uncertain status in both theoretical and historical phonology and, as was shown in chapter 3, I

believe that this is for very good reasons. The investigation of historical lenitions which has been undertaken here, in part accompanied by discussion of some synchronic lenitions, has shown that there is no formal unity among the group of 'lenition processes', but that they can nonetheless be seen to have certain properties in common, including the fact that, even though they are not caused by an the phonological environment in which they occur, they can be inhibited by it.

In this chapter, I first summarise the key claims that have been made in the thesis, in section 5.1. After this, I focus on what I believe to be the philological contributions of the work in section 5.2, and then on what I see as the key phonological contributions, in section 5.3. The next section, 5.4, brings together the empirical predictions that are made by the proposals in the thesis, and the final section, 5.5, discusses certain avenues for future research and concludes this conclusion.

5.1 Summary of analyses and claims

In chapter 1, as well as pointing forward to the areas of phonology, philology and historical linguistics that are discussed in subsequent chapters, I presented an approach to historical phonology which sees the concerns of 'naturalness' to be important in constraining and explaining the types of phonological process that can be endogenously innovated. All of the processes that have been discussed in detail in the thesis were endogenously innovated, apart from those discussed in section 4.2, where it was shown how aspects of exogeny can play a role in history phonology, too (as is well recognised in philological work). The lenition processes that have been discussed here were all innovated due to either phonetic or phonological factors which are inherent in the system and this makes them natural.

We also saw in chapter 1 that it is not straightforward to claim that such factors 'explain' the innovation of the lenitions, but this is the normal terminological practice in historical phonology, probably rightly so, given the nature of the discipline. In chapter 2, I introduced the notion that synchronic phonology should be 'minimalist' and this is like naturalness as it is a constraining factor on what is possible in phonology, but it is not the same as naturalness, which simply conditions process innovation. Considerations of minimalism mean that phonological derivations cannot venture too far from the surface, although they doubtless do, minimally, at times.

Chapter 1 also included the claim that the environmental conditioning of past phonological processes becomes fossilised into the underlying representations of a language, once a process

have been lexicalised. The notion of minimalism in phonology means that certain phonological processes become lexicalised quite quickly into the underlying phonology, although, if they are well motivated and transparent, phonological processes can survive for centuries. It seems likely that the synchronic Spanish lenition described in section 3.2.1.2 has been active in the phonology of that language for centuries, and the Liverpool English lenitions have probably been part of the variety since the mid nineteenth century. We might well assume that the approximantisation proposed dialectically for Germanic as part of the GCS₂ also survived for quite a while in the dialects into which it was innovated; indeed, it may well have started only minimally while Germanic was still a relatively coherent language and developed extensively only once the individual languages were relatively separate. As we saw in section 4.1.1, the process was certainly inhibited differently in different Germanic languages (perhaps most in the ancestor dialects of German and least in the ancestor dialects of Dutch).

Much of the philological discussion in chapter 2 has withstood the onslaught of phonological theory, although the symbolic representations of the processes concerned were quite different by the end of chapter 4 from what they were in chapter 2. The philological foundation has proved indispensable and, while I hope to have contributed slightly to the interpretation of philological material, and to have improved our understanding of what occurred in the GCS₂ and the HGCS₁, for example, much of the reanalysis which occurs in chapter 4 is reinterpretation. This process of reinterpretation is important as it helps to understand the causation of the processes involved which allows us to better explain them (even if we accept the caveats on the use of 'cause' and 'explain' which were discussed in chapter 1). The Glottalic Theory of IE stop consonantism was introduced at the end of chapter 2 and some of the analyses in chapter 4 work with the reconstructions which were proposed by Gamkrelidze & Ivanov (1973), and are similar to those of Hopper (1973) and others. If the stops of IE are given their traditional values, then some of the analyses in section 4.1.1 will need a little reworking, but even then, the essentials of them will remain the same.

The discussion in chapter 3 covered a quite wide range of issues. The basics of a model of phonology were introduced, relying chiefly on work which has been carried out in autosegmental non-linear frameworks, such as Government Phonology and Dependency Phonology, but which also has clear connections, I have argued, with work in other models, such as Feature Geometry and Articulatory Phonology, for example. The main focus in the first half of chapter 3 (section 3.1) was on melodic representation, but a short discussion of prosody was included, showing how the melodic material is tied into words (through the media

of the segment, root node and timing slot). The discussion of prosody proved to be important in the final understanding of the HGCS₁.

The model of melodic representation used here makes use of privative elements which are most like the units used in Government Phonology, but are far from incomparable with the units of other phonological traditions. The main focus in terms of melody was on laryngeal specifications. We saw in section 3.1.3.1 that these have often not been clearly understood before. It was shown that some key recent work in phonology has succeeded in formalising the philological insight behind the notion of the 'fortis/lenis' distinction and this led us to recognise the state of affairs which I labelled 'laryngeal realism'. This is simply the fact that different languages make use of different laryngeal elements, just as they do in other aspects of melody. The recognition of this provided a spur in the discussion of lenition in section 3.2, where it played an important role in the understanding of lenition processes that was developed there. The identification of laryngeal realism means that the laryngeal history of many languages has been miswritten and this recognition informed the analyses which were proposed for the Germanic data in chapter 4.

The rest of the second half of chapter 3 (section 3.2), involved a discussion of some other attested cases of lenition processes, and this played a vital role in the general discussion of lenition. The processes discussed there included examples from quite a range of languages, separated from each other in terms of history, geography and genetics. The generalised discussion of lenition which followed this showed firstly that previous approaches to lenition have not succeeded in unifying the processes involved as one process-type, although some of the previous approaches were seen to be more successful than others. The understanding of lenition which I proposed in section 3.2.2.3 was not a unificatory one. It was proposed that the processes simply have in common the fact that they are innovated 'spontaneously' as weakly unconditioned processes. They are not caused by their environment, but can be partially inhibited by it. The kind of definition which was adopted in that section was defended on philosophical grounds. The final part of chapter 3 investigated aspects of the interaction between the disparate lenition processes and the phonological environments in which the segments that they affect occur. I proposed that this should best be understood as a situation where prosodic and melodic factors can inhibit lenitions, and I developed an explanation for melodic lenition inhibition which attempts to account for 'inalterability' effects in a principled way. The strength that some segments show in the face of lenition processes was derived overtly from autosegmental element sharing. This notion that sharing gives segments strength was shown to apply to quite a wide range of circumstances.

Much of the discussion that took place in chapter 4 has already been summarised. The point of the chapter was partly to illustrate the result of the combination of the points which had been discussed previously in the thesis and, partly, the analyses are meant as contributions to philology and phonology in their own right. It was shown that many of the processes discussed there were related to the fact that Proto-Germanic was a *|spread|* language. The *|spread|* specification led to widespread affrication and spirantisation twice in the history of the Germanic languages (as it has done again in Liverpool English), as its loss has led to processes which have been interpreted as voicings, but can, in fact, be seen to be delaryngealisations. The remainder of chapter 4 (section 4.2) dealt with certain outstanding issues in Germanic laryngeal history, and showed how exogeny can affect languages.

In the remainder of this chapter, I draw out some of the specifics of the points summarised above and briefly examine some objections that could be raised to them.

5.2 Philological contributions

One of the aims of this thesis was to contribute to philology, at least in part by illustrating how aspects of phonological theory can have implications for the data which is discussed in philological work. I have also made certain specific claims and analyses which, I believe, contribute to philological concerns.

Firstly, I proposed that the letter which is written in early German texts as *⟨z⟩* (which is how it is often given in historical Grammars of the type used as the basis for chapter 2) or as *⟨z⟩*, should be interpreted as */θ/*. This was on the basis of the analoguehood of the lenition which produced this segment as one of its outputs to the lenition found in present-day Liverpool English (and described in some detail in section 3.2.1.8). It seems reasonable that such a reconstruction is best based on an attested linguistic state which can be seen to be extremely similar in many other respects. Also in connection with the HGCS₁, I showed that the traditional formulation cannot be correct because the spirantisation process could not produce geminate fricatives. I proposed an alternative scenario which captures the facts just as well, by assuming that a later gemination occurred in phonologically possible environments.

It is possible also that the discussion here has provided evidence for the interpretation of (at least one of) the IE laryngeals. This was provided by the process described here as ‘Indo-Aryan *TA* formation’, which was shown in section 4.1.1 to have been generalised in Germanic.

This indicates that the laryngeal involved was a glottal fricative, because it is the glottal fricative which is composed only of |spread|.

The discussion of the possibilities in lenition, along with the recognition of the patterns of lenition inhibition, has provided important back-up to the ‘minority’ interpretation of the GCS₂. It was shown in section 4.1.1 that the traditional analysis is unparsimonious and highly unlikely as it goes against everything that seems to be observable in lenition inhibition. The alternative scenario proposed in chapter 4 is equally compatible with the data and with the insights of phonology. In the next section, I explain how I believe this thesis has contributed to those insights.

5.3 Linguistic contributions

This thesis has also aimed to contribute to our understanding of phonology, largely through the focus on the notion of lenition. Much of the key data for this is historical, and this illustrates the implications that a proper understanding of diachrony can have for theoretical models of phonology. The recognition of laryngeal realism has relied on the work of others but the application of it to help explain lenition processes is novel and, as we have seen, has some quite important implications for phonological theory. We saw in section 4.1.4 that it is the recognition of laryngeal realism that allows us to rescue Hyman/Jakobson/Greenberg’s phonological universal about what is a possible phonological process and possible phonological system from what would otherwise seem to be disconfirmatory counterevidence.

We have also seen in this thesis that there are grounds to recognise phonological ‘molecules’, as well as ‘particles’ and ‘atoms’. The metaphor here can help us to understand the patterns that are found in lenition inhibition. The phonological particle is the smallest unit - the element (or feature). These combine to create atoms - segments - which are vital building blocks in formation of matter - words. But atoms can also combine by bonding with other atoms to form molecules - partial (or full) geminates. It is these phonological molecules that can give the segments which make them up the strength to resist lenition. I turn now to the detailed discussion of lenition in a dedicated subsection.

5.3.1 Obstruent lenition

In the initial discussion of lenition in section 1.4, it was noted that it is just over 100 years since the term was invented (in Thurneysen 1898). As we saw especially in section 3.2.2, the

term has caught the imagination of phonologists, in both synchronic and diachronic discussion and ideas connected with it have played a role in the development of phonological theories. The conclusions that I draw in section 3.2.2.3 may thus seem disappointing, but I believe that, while we are justified in treating the processes as a group in some sense, this might almost be thought to be because of what they are not, rather than what they are. Lenitions are not assimilations or dissimilations or epentheses, they are weakly unconditioned segmental processes.

The consideration of obstruent lenition in this thesis has shown that it can involve the loss of elemental material in various ways, and that it can also involve acoustic reanalysis. It is likely that if other processes which have also been described as lenition are considered, then the number of process-types involved will increase. It can indeed, though, still be helpful to group these processes together, partly for the understanding of phonological change.

It may be that lenition is best thought of by fitting segments on a lenition trajectory, as long as such trajectories are not accorded any formal existence in a theory. We can generalise further about some of the properties that lenitions tend to show, but this is generally because we have defined the category 'lenition' so as to include these properties: (i) they do not affect the major place of articulation of a segment; (ii) they often affect series of segments; (iii) they tend to be inhibited according to similar prosodic and melodic environments; (iv) they can be fitted on a trajectory built up on Vennemann's personal communication in Hyman (1975) which was quoted in section 1.4 and which relies on the observations of historical 'paths towards deletion'.

As humans, we tend to perceive the world in terms of discrete categories and concepts. If we have a concept 'lenition' then we can perceive process types as being examples of it and we will want to identify commonalities among the members of the class. But it is not always the case that a class of items consists of entities that are ontologically coherent. For example, the class 'wings' includes birds' wings and bats' wings, which both evolved in rather different ways and function quite differently, but serve a similar purpose and clearly have certain properties which can allow us to group them together; we might even include the wings of an aeroplane in the class. Equally, the class of 'heavenly body' includes stars, planets and satellites, which are very different kinds of things, but nonetheless have certain things in common and to the observer it makes sense to say that they belong to a class. It seems that lenition is the same kind of class.

We can still recognise generalisations in lenition processes, certainly in terms of the patterning of the individual processes involved. Indeed, as I have explained, particularly in

section 3.2.2.3, there are certain generalisations that can be made about the processes but which have not been properly made before, both in terms of the factors which lead to particular processes being innovated and in terms of their inhibition.

Most lenitions are inhibited, even if only slightly. As we have seen, where there is only minimal inhibition, the tends to be melodic inhibition and will thus typically involve full geminates or partial place or laryngeal geminates. Many lenitions exhibit prosodic inhibition as well, though, and this can play an important role in the full description of a lenition. Although we have not focussed on it in this thesis, it seems clear that initial environments can be inhibitory. In terms of precisely which inhibitory factors are active in any particular lenition, we have to admit that chance plays a role. Just as we cannot predict precisely when a process will be innovated, we also cannot predict precisely what patterning a process will have, should an innovation occur. We can make predictions about what will not occur, however, and I turn to these in the next section.

5.4 Empirical predictions

In section 1.3.1, we saw that negative strong predictions are in principle falsifiable, as are all types of absolute predictions, whereas positive strong predictions, just like weak predictions, do not really make empirical claims as there is no cut-off point when we could recognise that they have been proven or disproven. With negative strong predictions (ie, that something cannot occur, although this is not tied to a particular time and place), there is a clear cut-off point. If we find one clear counterexample, then the prediction is disproven. Several claims made in this thesis have the status of empirical predictions. In this section, I explain how this is so.

In section 3.2.2.3 it was predicted from the explanation that was offered for spirantisation that true spirantisation could only occur through affrication, and because affrication of this spontaneous type relies on the presence of $|\text{spread}|$, it was predicted that spirantisation can only occur in $|\text{spread}|$ languages. This is a strong prediction and the negative version which it implies is that a process such as $k^{\circ} \Rightarrow x^{\circ}$ cannot be innovated in diachronic phonology. It was also predicted in section 3.2.2.3 that $g \Rightarrow \gamma$ cannot be innovated, where $/\gamma/$ is interpreted as the true fricative that the symbol stands for in the IPA. A further prediction is that only $|\text{spread}|$ segments can debuccalise to $/h/$.

These are all empirical hypotheses which could be falsified by the discovery of a lenition which has been naturally innovated and which has the properties which are proposed to be impossible. If such processes can truly be found and cannot be accounted for in other ways, the particular hypothesis about the diachronic effects of being a |spread| language would be disproven. This would not disprove the entire laryngeal realist position, of course, but it would disprove certain aspects of the way in which I have implemented it.¹

One apparent such counterexample is the pattern of lenition found in Tuscan Italian, often referred to as the *gorgia toscana* (see Izzo 1972, Vincent 1988, Maiden 1995, Kirchner 1998). This has been reported to involve the ‘spirantisation of voiceless stops’, and, importantly, Italian, as a Romance language, might be expected to be a |voice| language and so there would be no |spread| in the *T*s which could lead to such spirantisation. If this were truly the case, then it would be precisely the kind of situation that would falsify my claim. However, there are several points which disarm it and show that, in fact, the *gorgia* is compatible with what I have claimed. There are in fact two separate ways in which the *gorgia* can be shown to be compatible with my proposals. The first is that there are reports of aspiration for Tuscan Italian. For example, Vincent (1998) writes that

Another kind of intervocalic weakening is to be found in Tuscany where Latin /p t k/ develop into aspirates or fricatives - the so-called *gorgia toscana* (‘Tuscan throat’) - hence such common regional pronunciations as [kəhəhola] for *Coca Cola*, [stat^ho] or [staθo] for *stato* ‘been’, [pip^ha] or [piɸa] for *pipa* ‘pipe’

Vincent (1988, 287)

The presence of aspiration in this variety has been linked to an Etruscan substrate in the area where the *gorgia* occurs (see, for example Izzo 1972, Maiden 1995) and it may be that this is the ultimate original of the |spread| specification. Whatever its origin, if it is there, then the possibility is clear for it to be phonologically active and hence, diachronically to effect such spirantisation. Even if this is not accepted, there is further evidence to suggest that the Tuscan *gorgia* is not counterevidential to my proposal. The other way in which it could be shown to be compatible is if this Tuscan lenition was not spirantisation but approximantisation. This should not be assumed lightly, or else the hypothesis could become worthless, but it seems that detailed phonetic investigations may back this up. For example, Kirchner (1998) writes of the *gorgia* that “[i]n intervocalic position, the voiceless stops /p,t,k/ are obligatorily spirantized,

¹ The |spread|-language/|voice|-language distinction itself is empirical in other ways. It could be falsified if, say a language is found with copious assimilation to voice, but with aspiration of *T*s and an utter lack of voicing in *Ms*.

typically to approximants [ϕ, θ, ʃ]" (1998, 253). While it may seem odd that Kirchner describes this as spirantisation, it seems that the outputs may well be approximants, which could straightforwardly be accounted for by the loss of [occlusion].

There are thus two possible explanations for the Tuscan *gorgia*, and both are compatible with the predication made in this thesis. If there truly is spirantisation, then we may attribute it to a [spread] specification, which may have been imported into Tuscan Italian exogenously, in a way similar to those described in section 4.2, or it may even be that the lenition does not feature spirantisation at all. Where other similar cases to this come to light, however, the potential for falsification is still apparent, if they cannot be accounted for in other ways, as is the case for the Tuscan *gorgia*.

The other main source of empirical predictions in this thesis lies in the proposals which were developed to account for melodic lenition inhibition. The key proposal there was that the autosegmental sharing of elements can bond segments together and give them the 'strength' to resist the lenition process. It is clear that this is not absolute strength, as lenitions can occur even to partial and full geminates. It is also the case, as we have seen, that prosodic inhibition can fail, so the initial segments are lenited. We saw in the discussion of the GCS₁ that partial place geminates of the type [N__] are not always inhibitory, although in that case lenition was inhibited in the partial laryngeal geminates.

This lack of absolute predictive power naturally weakens the explanatory power of the proposal, although this will affect any account which seeks to explain these facts, because the facts themselves are disparate. The proposal does still make falsifiable predictions however. It would be falsified if a language were found where a lenition process occurs in partial laryngeal geminates, for example, or in partial place geminates, but not in intervocalic stops, or in those in coda position, because these are predicted to be non-inhibitory positions. Also, it would be falsified if, in a lenition, stops in such clusters as [lk] or [rp] were not lenited, where as those in [nt] and [ŋk] were.

There are certain types of process which might be thought of as counterevidence to the proposal, such as the fate of /sk^h/ clusters in many varieties of Germanic, but as they stand, they do not seem to falsify the key claims. The cluster /sk^h/ has been lost in English, Dutch and German, quite possibly as independent developments (see Nielsen 1980 for some discussion of this point). In English and German, the cluster has fused to become /ʃ^h/, and in most varieties of Dutch, the cluster is now pronounced /sx/. This might be thought of as a counterexample as it seems to show a lenition which only occurs in a position which is predicted to be strong.

However, it is far from clear that this process should be seen as a case of lenition. The environment is so restricted in this situation that the process would be a very unusual type of lenition as it would violate all generalisations about prosodic inhibition as well, because it did not occur intervocalically or finally. It is notable also that the process only affected /k/, whereas, as we have seen, lenitions often affect series of segments

In the English and German case, the process seems better described as a case of fusion, or coalescence, which can be conceived of as a case of mutual assimilation, indeed, such a process could actually be seen as the tightening of phonological bonds, as the two segments are drawn into one. For these reasons, it does not seem clear that phonologically isolated cases such as this should count as counterevidence to the wider proposals for melodic inhibition. Evidence of the type outlined above, however, would be. The proposals outlined in this section thus have some clear falsifiable predictive power. While they have not been falsified in this section, they could still in principle be falsified and thus may be seen, with certain caveats, as explanatory.

5.5 Future research and final conclusion

The issues just addressed require further consideration. If such empirical predictions are made, then part of the task of future research is to test them, partly in the way which has already been outlined in the last section. If they turn out to be true, then this testing process will increase our understanding of the data which is investigated with the aid of them. If they turn out to be false, then they will still have served a purpose, and the phonological room for hypothesising will be reduced.

Various other aspects of the thesis open up the way for future research. The assumption of laryngeal realism means that the laryngeal phonology of many languages may have been miswritten and it will be important to consider the ideas discussed here in relation to languages other than Germanic. There are also still many other diachronic obstruent occurrences in the history of Germanic languages, and this data is readily available in the philological Handbooks and Grammars, as described in chapters 1 and 2 of this thesis. It is likely that a consideration of such data will help to further increase our understanding of phonology in general and of those pieces of data in particular, in the same way that the investigations in this thesis have done. It would likely also be potentially fruitful to consider other types of phonological processes which have been described as cases of lenition in the light of the discussion that has occurred in this thesis

Certain ideas which have been discussed in the thesis have been left a little vague in the discussion here, and future research will help to firm these up. One of these is the notion of the diachronic quantum. It is not yet clear what counts as a possible quantum, but it is to be hoped that future research will consider the issue. The mechanism behind prosodic lenition inhibition was also left unformulated, and this is a further avenue for future research, as is a general refining of the notion of melodic inhibition and the 'strength-through-sharing' approach which has been developed here to account for it.

A final issue which has been left unaddressed here is the precise place of articulation that is taken on by the outputs of spirantisation and maybe of approximantisation, too. While it is true that lenition does not alter the major place of articulation of a segment, there may be a certain change 'within a major articulator'. In section 3.1.5, these distinctions were accounted for by a difference in headedness among the elements and it may be that there are certain generalisations which can be made in connection with this point. There is also scope for a reformulation of some of the elements that were used in this thesis, as was explained when they were introduced. This is particularly likely to be the case for the place elements.

I hope that the research which is to follow from that which is reported on here will continue to illustrate the connections that exist between philology and phonology, between a historical approach and a theoretical approach. While they spent much of their development as disciplines as mutually influencing fields of study, it seems sometimes that they now can be too easily separated. While some mutual implications have been explored in this thesis, each discipline still holds many more, which are as yet unexplored, for the other.

References

- Albrecht, K. (1881) *Die Leipziger Mundart; Grammatik und Wörterbuch der Leipziger Volksprache*. Leipzig: Arnoldsche Buchhandlung.
- Alexander, G (1983) *Fortis and Lenis in Germanic*. New York & Berne: Peter Lang.
- Aitchison, J. (1987) 'The Language Lifegame: Prediction, Explanation and Linguistic change'. In: Koopman, van der Leek, Fischer & Eaton (1987).
- Ammon, U. & Loewer, U. (1977) *Schwäbisch. Dialekt/Hochsprache kontrastiv*. Volume 4. Düsseldorf: Padagogischer Verlag Schwann.
- Anderson, J. & Durand, J. (1986) 'Dependency Phonology' In: Durand (1986).
- Anderson, J. & Durand, J. (eds) (1987) *Explorations in Dependency Phonology*. Dordrecht: Foris.
- Anderson, J. & Ewen, C. (1987) *Principles of Dependency Phonology*. Cambridge: CUP
- Anderson, J. & Jones, C. (1974) 'Three theses concerning phonological representations' *Journal of Linguistics* 10, 1-26.
- Archangeli, D. & Suzuki, K. (1997) 'The Yokuts challenge'. In: Roca, I. (ed) *Derivations and Constraints in Phonology*. Oxford: Clarendon.
- Arens, H. (1969) *Sprachwissenschaft. Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*. Second edition. Freiburg: Karl Alber.
- Arlotto, A. (1972) *Introduction to Historical Linguistics*. Boston: Houghton Mifflin.
- Azra, J. (1994) Paper presented at the *International Conference on of Historical Linguistics*, University of Manchester.
- Bach, E. & Harms, R. (1972) 'How do languages get crazy rules?' In: Stockwell, P. & Macaulay, R. (eds) *Linguistic Change and Generative Theory*. Bloomington: Indiana University Press.
- Ball, M. & N. Müller (1992) *Mutation in Welsh*. London: Routledge.
- Bammesberger, A. (1988) 'Voice opposition in the system of Old English spirants' In: Kastovsky, D. & Bauer, G. (eds) *Luick Revisited*. Tübingen: Gunter Narr.
- Bammesberger, A. (1992). 'The place of English in Germanic and Indo-European' In: Hogg (1992b) .

- Bammesberger, A. & Wollman, A. (eds) *Britain 400-600: Language and History*. Heidelberg: Carl Winter.
- Barbour, S. & Stevenson, P. (1990) *Variation in German*. Cambridge: CUP.
- Bartholomae, C. (1894) 'Vorgeschichte der iranischen Sprachen' In: Geiger & Kuhn (1895-1901).
- Bartholomae, C. (1895) 'Awestsprache und Altpersisch' In: Geiger & Kuhn (1895-1901).
- Bauer, L. (1988) 'What is lenition?' *Journal of Linguistics* 24, 381-392.
- Beekes, R. (1995) *Comparative Indo-European Linguistics*. Amsterdam: John Benjamins.
- Beekes, R. (1997) 'Historical phonology of Iranian' *Journal of Indo-European Studies* 25, 1-26.
- Bennet, J. (1955) 'The southern English development of Germanic *[f s þ]' *Language* 31, 367-371.
- Bergmann, G. (1990) 'Upper Saxon' In: Russ (1990).
- Bermúdez-Otero, R. (1999) *Constraint Interaction in Language Change: Quantity in English and German*. Unpublished PhD, University of Manchester.
- Bermúdez-Otero, R. (2000) 'Voicing and continuancy in Catalan; a nonvacuous Duke-of-York gambit and Richness-of-the-Base paradox'. University of Manchester, Manuscript.
- Bickmore, L. (1995) 'Accounting for compensatory lengthening in the CV and moraic frameworks' In: Durand & Katamba (1995).
- Bloch-Rozmej, A. (1995) 'Nasal Lenition and the Theory of phonological Government.' In: Gussman, E. (ed) *Licensing in Syntax and Phonology*. Lublin: Wydawnictwo Folium.
- Bloomfield, L. (1933) *Language*. New York: Holt.
- Bock, G. (1965) *Die Mundart von Waldau bei Schleusingen*. Cologne & Graz: Böhlau Verlag.
- Bomhard, A. (1986) 'The aspirated stops of Proto-Indo-European' *Diachronica* 3, 67-79.
- Booij, G. (1995) *The Phonology of Dutch*. Cambridge: CUP.
- Bopp, F. (1816) *Über das conjugationssystem der Sanskritsprache*. Reprinted (1975) Hildesheim: Olms.

- Braun, A. (1988) *Zum Merkmal "Fortis/Lenis"; Phonologische Betrachtungen und instrumentalphonetische Untersuchungen an einem mittelhessischen Dialekt*. Zeitschrift für Dialektologie und Linguistik Beiheft 55. Stuttgart: Franz Steiner.
- Braune, W.. (1880) *Gotische Grammatik*. Sammlung kurzer Grammatiken germanischer Dialekte. Halle: Niemeyer.
- Braune, W. (1891) *Althochdeutsche Grammatik*. 2nd edition. Sammlung kurzer Grammatiken germanischer Dialekte. Halle: Niemeyer.
- Braune, W. & Ebbinghaus, E. (1981) *Gotische Grammatik*. 19th Edition. Sammlung kurzer Grammatiken germanischer Dialekte. Tübingen: Niemeyer.
- Braune, W. & Eggers (1987) *Althochdeutsche Grammatik*. 14th Edition. Sammlung kurzer Grammatiken germanischer Dialekte. Tübingen: Niemeyer.
- Bree, C. van (1987) *Historische Grammatica van het Nederlands*. Dordrecht: Foris.
- Broadbent, J. (1999) 'The particle representation of consonant place' *Roehampton Institute London Working Papers in Linguistics* 1, 50-96.
- Brockhaus, W. (1995a) *Final Devoicing in the Phonology of German*. Tübingen: Niemeyer.
- Brockhaus, W. (1995b) 'Skeletal and suprasegmental structure within Government Phonology' In: Durand & Katamba (1995).
- Bromberger, S. & Halle, M. (1989) 'Why phonology is different' *Linguistic Inquiry* 20, 51-70.
- Browman, P. & Goldstein, L. (1989) 'Articulatory gestures as phonological units' *Phonology* 6, 219-52.
- Browman, P. & Goldstein, L. (1992) 'Articulatory Phonology: an overview' *Phonetica* 49, 155-180.
- Brugge, E. (1944) *Vokalismus der Mundart von Emmerstedt: mit Beiträgen zur Dialektgeographie des östlichen Ostfalen*. Lund: Lunder Germanischische Forschungen.
- Brugmann, K. & Delbrück (1886-1900) *Grundriss der vergleichende Grammatik der indogermanischen Sprachen*. Five volumes. Strasburg: Trübner.
- Brunner, K. (1965) *Altenglische Grammatik; nach der angelsächsischen Grammatik von Eduard Sievers neubearbeitet*. 3rd edition. Tübingen: Niemeyer.
- Buck, C. (1933) *Comparative Grammar of Greek and Latin*. Chicago: University of Chicago Press.

- Burton-Roberts, N. (1994) 'Ambiguity, Sentence and Utterance: A Representational Approach' *The Transactions of the Philological Society* 92, 179-212.
- Burton-Roberts, N. (2000) 'Where and what is phonology? A representational perspective' In: Burton-Roberts, Carr, & Docherty (2000).
- Burton-Roberts, N. & Carr, P. (1999) 'On speech and natural language' *Language Sciences* 21, 371-406.
- Burton-Roberts, N., Carr, P. & Docherty, G. (eds) (2000) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: OUP.
- Butterfield, J. et al (1997) *Collins Spanish Dictionary*. Glasgow: Harpercollins.
- Bynon, T. (1977) *Historical Linguistics*. Cambridge: CUP.
- Campbell, L. (1998) *Historical Linguistics; an Introduction*. Edinburgh: EUP.
- Carr, P. (1993) *Phonology*. London: Macmillan.
- Carr, P. (2000) 'Scientific realism, sociophonetic variation, and innate endowments in phonology' In: Burton-Robert, Carr & Docherty (2000).
- Catford, J. (1988) *A Practical Introduction to Phonetics*. Oxford: Clarendon.
- Caudmont, J. (2001) 'In memoriam: André Martinet' *Linguistische Berichte* 185, 3-12.
- Charette, M. (1991) *Conditions on Phonological Government*. Cambridge: CUP.
- Chomsky, N. (1986) *Knowledge of Language*. New York: Praeger.
- Chomsky, N & Halle, M. (1968) *The Sound Pattern of English*. New York: New York: Harper & Row.
- Churma, D (1988) 'On "On Geminates": an alternative approach to inalterability' SUNY at Buffalo, Manuscript.
- Clements, N. (1985) 'The geometry of phonological features' *Phonology* 2, 225-252.
- Clements, G. & Hume, E. (1995) 'The internal organization of speech sounds' In: Goldsmith (1995).
- Coetsem, F. van (1970) 'Zur Entwicklung der germanischen Grundschrache' In: Schmitt, E. (ed) *Kurzer Grundriß der germanischen Philologie*. Berlin: de Gruyter.

- Cohen, A., Ebeling, C., Fokkema, K., Holk, A. van (1972) *Fonologie van het Nederlands en het Fries*. The Hague: Martinus Nijhoff.
- Cohen, V. (1971) 'Foleyology' CLS 7.
- Collinge, N. (1985) *The Laws of Indo-European*. Amsterdam: John Benjamins.
- Collins, B. & Mees, I. (1981) *The Sounds of English and Dutch*. The Hague: Leiden University Press.
- Crowley, T. (1992) *An Introduction to Historical Linguistics* Auckland: OUP.
- Cutler, A. (1992) 'Phonology and the segment' In: Docherty & Ladd (1992).
- Davenport, M. & Hannahs, SJ (1998) *Introducing Phonetics and Phonology*. London: Arnold.
- Davies, A. (1997) *History of Linguistics. Volume 4: Nineteenth-Century Linguistics*. London: Longman..
- Davis, G. & Iverson, G. (1995) 'Segment Organization in the High German Consonant Shift' *American Journal of Germanic Linguistics and Literatures* 7, 111-127.
- Davis, G., Iverson, G. & Salmons, J. (1999) 'Peripherality and markedness in the spread of the High German Consonant Shift' *Beiträge zur Geschichte der deutschen Sprache und Literatur* 121, 177-200.
- Deutscher, G. (2000) 'Akkadian Language Tutorial' *Linguistics Association of Great Britain Autumn Meeting*. University of Durham.
- Dienes, P. & Szigetvári, P. (1999) 'Repartitioning the skeleton: VC phonology' Manuscript.
- Dixon, R (1997) *The Rise and Fall of Languages*. Cambridge: CUP.
- Dobson, E. (1968) *English Pronunciation 1500-1700. Volume 2: Phonology*. Second edition. Oxford: Clarendon.
- Docherty, G. & Ladd, R. (eds) (1992) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: CUP.
- Docherty, G., Foulkes, P., Milroy, J., Milroy, L. & Walshaw, D. (1997). Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics* 33: 275-310.
- Doke, C. (1954) *The Southern Bantu Languages*. London: OUP.
- Donegan, P. & Stampe, D. (1979) 'The study of Natural Phonology' In: Dinnsen, D. (ed) *Current Approaches to Phonological Theory*. Bloomington: Indiana University Press.

- Dosuna, J. (1996) 'Can weakening processes start in initial position? The case of aspiration of /s/ and /f/' In: Hurch, B. & Rhodes, R. (eds) *Natural Phonology: the State of the Art*. Berlin: Mouton de Gruyter.
- Dresher, B. E. (1993) Review of Hogg (1992) *Phonology* 10, 146-157
- Dressler, W. (1985) *Morphonology: the dynamics of derivation*. Ann Arbor: Karoma.
- Duckworth, M., Allen, G., Hardcastle, W. & Ball, M. (1990) 'Extensions to the International Phonetic Alphabet for the transcription of atypical speech'. *Clinical Linguistics and Phonetics* 4, 273-280.
- Durand, J. (ed) (1986) *Dependency and Non-linear Phonology*. London: Croom Helm.
- Durand, J. (2000) 'The new IPA handbook and the British tradition' Paper presented at the conference on *Linguistics and the English Language*. Université de Toulouse le Mirail.
- Durand, J. & Katamba, F. (eds) (1995) *Frontiers in Phonology*. London: Longman.
- Eaton, R. & Koopman, W. (1987) 'Introduction' In: Koopman, W., Leek, F, van der, Fischer, O. & Eaton, R. *Explanation and Linguistic Change*. Amsterdam: John Benjamins.
- Ellis, A. (1869) *Early English Pronunciation*. London: The Philological Society.
- Elmedlaoui, M. (1993) 'Gemination and spirantization in Hebrew, Berber and Tigrinya: a "Fortis-Lenis Module" analysis.' *Linguistica Communicatio* 5, 121-176.
- Emonds, J. (1972) 'A reformulation of Grimm's Law' In: Brame, M. K. (ed.) *Contributions to Generative Phonology*. Austin: University of Texas Press.
- Escure, G. (1975) *Weakening and Deletion Processes in Language Change*. Unpublished PhD, Indiana University.
- Escure, G. (1977) 'Hierarchies and phonological weakening' *Lingua* 43, 55-64.
- Ewen, C. (1995) 'Dependency relations in phonology' In: Goldsmith, J. (ed) *The Handbook of Phonological Theory*. Oxford: Blackwell.
- Ewen, C. & Hulst, H. van der (2001) *The Phonological Structure of Words*. Cambridge: CUP.
- Fallon, P. (1998) *The Synchronic and Diachronic Phonology of Ejectives*. Unpublished PhD, Ohio State University.
- Fischer-Jørgensen, E. (1975) *Trends in Phonological Theory*. Copenhagen: Akademisk Forlag.

- Fisiak, J. (1984) 'The voicing of Initial fricatives in Middle English' *Studia Anglica Posnaniensia* 17, 3-23.
- Fisiak, J. (ed) (1997) *Linguistic Reconstruction and Typology*. Berlin: Mouton de Gruyter.
- Fónagy, I (1963) *Die Metaphern in der Phonetik*. The Hague: Mouton.
- Hague. Foley, J. (1977) *Foundations of Theoretical Phonology*. Cambridge: CUP.
- Foley, J. (1981) 'Reply to Smith' In: Goyvaerts (1981).
- Fourquet, J. (1948) *Les Mutations Consonantiques du Germanique: Essai de Position des Problèmes*. Paris: Belles Lettres.
- Foulkes, P. (1993) *Theoretical Implications of the /p/ > /f/ > /h/ Change*. Unpublished PhD. Cambridge University.
- Foulkes, P. (1997) 'Historical laboratory phonology - investigating /p/ > /f/ > /h/ changes' *Language and Speech* 40, 249-276.
- Foulkes, P. & Docherty, G. (eds) (1999) *Urban Voices: Accent Studies in the British Isles*. London: Arnold.
- Fox, A. (1995) *Linguistic Reconstruction: an Introduction to Theory and Method*. Oxford: OUP.
- Fried, E. (1978) *100 Gedichte ohne Vaterland*. Paperback edition 1993. Frankfurt am Main: Fischer.
- Fudge, E. (1969) 'Syllables' *Journal of Linguistics* 5, 253-287.
- Gamkrelidze, T. & Ivanov, V. (1973) 'Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse'. *Phonetica* 27, 150-156.
- Gamkrelidze, T. & Ivanov, V. (1984) *Indo-European and the Indo-Europeans*. Translated (1995) by J. Nichols. Berlin: Mouton de Gruyter.
- Geiger, W & Kuhn, E. (1895-1901) *Grundriss der iranischen Philologie*. Straßburg: Karl Trübner.
- Ghosh, R. (1995) 'Compensatory lengthening in a dialect of Bangla' *CIEFL Occasional Papers in Linguistics* 5. Hyderabad: CIEFL.
- Giegerich, H. (1992) *English Phonology: an Introduction*. Cambridge: CUP.

- Giles, H. & P. Trudgill (1978). Sociolinguistics and linguistic value judgements; correctness, adequacy and aesthetics. Reprinted in Trudgill, P. (1983) *On dialect; social and geographical perspectives*. Oxford: Blackwell. 201-225.
- Goblirsch, K. (1994) 'Fortis and Lenis in Standard German' *Leuvense Bijdragen* 83, 31-45.
- Goldsmith, J. (1976) *Autosegmental Phonology*. PhD, MIT. Published (1979) New York: Garland.
- Goldsmith, J. (1989) 'Licensing, inalterability, and harmonic rule application' *CLS* 25, 145-156.
- Goldsmith, J. (1990) *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- Goldsmith, J. (1995) *Handbook of Phonological Theory*. Oxford: Blackwell.
- Goldsmith, J. (1999) 'Introduction' In: Goldsmith, J. (ed) *Phonological Theory: the Essential Readings*. Oxford: Blackwell.
- Goldsmith, J. & Laks, B. (2000) 'Introduction' In: Goldsmith, J. & Laks, B. (eds) *Folia Linguistica* 34, Special Issue: *The History of Phonology in the Twentieth Century*, 1-10.
- Goyvaerts, D. (1975) *Present-Day Historical and Comparative Linguistics*. Ghent: Story-Scientia.
- Goyvaerts, D. (ed) *Phonology in the 1980s*. Ghent: E. Story-Scientia.
- Greenberg, J. (1966) 'Synchronic and diachronic universal in phonology' *Language* 42, 508-517.
- Grijzenhout, J. (1995). *Irish Consonant Mutations and Phonological Theory*. Utrecht: OTS.
- Grimm (1822-37) *Deutsche Grammatik*. 4 volumes. Göttingen: Dieterich'sche Buchhandlung.
- Guerssel, M. (1978) 'A condition on assimilation rules' *Linguistic Analysis* 3, 267-305.
- Guthrie, M. (1967-1971) *Comparative Bantu*. 4 volumes. Farnborough: Gregg International.
- Haider, H. (1985) 'The fallacy of typology - remarks on the PIE-stop system' *Lingua* 65, 1-27.
- Hale, M. & Reiss, C. (2000) 'Phonology as cognition' In: Burton-Roberts, Carr & Docherty (2000).
- Halle, M & Mohanen, K. (1985) 'Segmental phonology of Modern English' *Linguistic Inquiry* 16, 57-116.
- Halle, M. & Stevens, K. (1971) 'A note of laryngeal features' *Quarterly Progress Report, Research Laboratory of Electronics*. MIT, 198-213.

- Hammarström, G. (1978) 'Is linguistics a natural science?' *Lingua* 45, 16-31.
- Hammond, M. (1999) *The Phonology of English; a Prosodic Optimality-Theoretic Approach*. Oxford: OUP.
- Harris, James (1969) *Spanish Phonology*. Cambridge, Mass: MIT Press
- Harris, James (1984) 'La espirantización en castellano y su representación fonológica autosegmental' In: Bartra et al (eds) *Estudis Grammaticals*. Universitat Autònoma de Barcelona.
- Harris, John (1985) *Phonological Variation and Change*. Cambridge: CUP.
- Harris, John (1990) 'Segmental complexity and phonological government' *Phonology* 7, 255-300.
- Harris, John (1994) *English Sound Structure*. Oxford: Blackwell.
- Harris, John (1994b) 'Minimalist phonology' Talk given at the International conference on language in Ireland. University of Ulster at Jordanstown.
- Harris, John (1997) 'Licensing inheritance: an integrated theory of neutralisation' *Phonology* 14, 315-370.
- Harris, John & Lindsey, G. (1995) 'The elements of phonological representation' In: Durand & Katamba (1995).
- Harris, John & Kaye, J. (1990) 'A tale of two cities: London glottaling and New York City tapping' *The Linguistic Review* 7, 251-274.
- Harris, John, Watson, J. & Bates, S. (1999) 'Prosody and melody in vowel disorder' *Journal of Linguistics* 35, 489-526.
- Harris-Northall, R. (1990) *Weakening Processes in the History of Spanish Consonants*. London: Routledge.
- Haugen, E. (1976) *The Scandinavian Languages*. London: Faber and Faber.
- Hayes, B. (1986) Inalterability in CV phonology. *Language* 62, 321-351.
- Hickey, R. (1984) 'Coronal segments in Irish English' *Journal of Linguistics* 20, 233-250.
- Hickey, R. (1996). 'Lenition in Irish English' *Belfast Working Papers in Linguistics* 13, 173-193.
- Hickey, R. (1999) 'Dublin English: current changes and their motivation' In: Foulkes & Docherty (1999).

- Hind, K. (1996) 'Lenition in Articulatory Phonology' *Belfast Working Papers in Linguistics* 13, 194-221.
- Hock, H. (1986) *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- Hock, H. & Joseph, B. (1996) *Language History, Language Change and Language Relationship*. Berlin: Mouton de Gruyter.
- Hoegniswald, H. (1960) *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
- Hogg, R. (1992a) 'Introduction' In: Hogg (1992b)
- Hogg, R. (ed) (1992b) *The Cambridge Historical or the English Language*. Volume 1: *The beginnings to 1066*. Cambridge: CUP.
- Hogg, R. (1992c) *A Grammar of Old English*. Volume 1: *Phonology*. Oxford: Blackwell.
- Honeybone, P. (1995) 'Explanation, Government Phonology and the history of velarity in German' *Newcastle and Durham Working Papers in Linguistics* 3, 77-104.
- Honeybone, P. (1996) 'Is lenition naturally optimal?' *Newcastle & Durham Working Papers in Linguistics* 4, 101-125.
- Honeybone, P. (1999) 'I blame the Government.' [A discussion of Government Phonology and Yoshida, S., 1996, 'Phonological Government in Japanese']. *Language Sciences* 21, 177-221.
- Honeybone, P. (2000a) Review of Hammond (1999). *Journal of Linguistics* 36, 165-171.
- Honeybone, P. (2000b) 'German *ich-Laut* and *ach-Laut* - noch einmal, schon wieder, again' Paper presented at the conference of the Forum for Germanic Language Studies, University of Manchester & UMIST.
- Honeybone, P. (2000c) 'When does synchrony become diachrony? The case of German dorsal fricatives' Paper presented at the eighth Manchester Phonology Meeting, University of Manchester.
- Honeybone, P. (2000d) 'Lenition in Liverpool English' Paper presented at the conference on Linguistics and the English Language, University of Toulouse le Mirail.
- Honeybone, P. (2001) 'Lenition Inhibition in Liverpool English' *English Language and Linguistics* 5, 213-249.
- Hooper, J. (1976) *An Introduction to Natural Generative Phonology*. New York: Academic Press.

- Hopper, P. (1973) 'Glottalized and murmured occlusives in Indo-European' *Glossa* 7, 141-166.
- Hornstein, N. & Lightfoot, D. (eds) (1981) *Explanation in Linguistics*. London: Longman.
- Horrocks, G. (1997) *Greek: a History of the Language and its Speakers*. London: Longman.
- Houlihan, K. (1977) 'On aspiration and deaspiration processes' In: Eckman, F (ed) *Current Themes in Linguistics: Bilingualism, Experimental Linguistics and Language typologies*. Washington: Hemisphere.
- Hulst, H. van der (1995) 'Radical CV Phonology: the categorial gesture' In: Durand & Katamba (1995).
- Hulst, H. van der & Ritter, N. (1999a) 'Head-Driven Phonology' In: van der Hulst & Ritter (1999b).
- Hulst, H. van der & Ritter, N. (eds) (1999b) *The Syllable: Views and Facts*. Berlin: Mouton de Gruyter.
- Hulst, H. van der & Smith, N. (1982) 'An overview of autosegmental and metrical phonology' In: Hulst, H. van der & Smith, N. (eds) *The Structure of Phonological Representations*. Dordrecht: Foris.
- Hyman, L. (1970) 'How concrete is phonology?' *Language* 46, 58-76.
- Hyman, L. (1975) *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.
- Hyman, L. (1985) *A Theory of Phonological Weight*. Dordrecht: Foris.
- Hyman, L. (1999) 'Lenition and fortition in Bantu' Paper presented at the round table on the strong position at the inaugural conference of the GDR Phonologie, University of Nice.
- IPA (1999) *Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet*. Cambridge: CUP.
- Iverson, G. & Salmons, J. (1995) 'Aspiration and laryngeal representation in Germanic' *Phonology* 12, 369-396.
- Iverson, G. & Salmons, J. (1999a) 'Glottal Spreading Bias in Germanic' *Linguistische Berichte* 178, 135-151.
- Iverson, G. & Salmons, J. (1999b) 'Peripherality and markedness in the spread of the High German Consonant Shift' *Beiträge zur Geschichte der deutschen Sprache und Literatur* 121, 177-200.

- Izzo, H. (1972) *Tuscan & Etruscan; the problem of linguistic substratum influence in central Italy*. Toronto: University of Toronto Press.
- Jacobs, H. (1994). Lenition and Optimality Theory. *Proceedings of the LSRL XXIV*. [ROA-127, Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>].
- Jakobson, R. (1958) 'Typological studies and their contribution to historical comparative linguistics' *Proceedings of the 8th Congress of Linguists*. Reprinted (1971) in: *Selected Works: Phonological Studies*. The Hague: Mouton.
- Jakobson, R. (1976) *Six leçons sur le son et le sens*. Paris: Editions de Minuit.
- Jakobson, R., Fant, G. & Halle, M. (1952) *Preliminaries to Speech Analysis*. Cambridge, Mass: MIT Press.
- Jakobson, R. & Waugh, L. (1979) *The Sound Shape of Language*. Brighton: Harvester Press.
- Jeffers, R. & Lehistes, I. (1979) *Principles and Methods for Historical Linguistics*. Cambridge, Mass: MIT Press.
- Jespersen, O. (1897) 'Karl Verner; some personal recollections' Translated and reprinted in Jespersen, O. (1933) *Linguistica; Selected Papers in English, French and German*. Reprinted (1970). College Park: McGrath Publishing Company.
- Jespersen, O. (1922) *Language: its Nature, Development and Origin*. London: Allen and Unwin.
- Jessen, M. (1997). *The Phonetics and Phonology of the Tense and Lax Obstruents in German*. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart. Published in a revised form as Jessen (1999).
- Jessen, M. (1999) *Phonetics and Phonology of Tense and Lax Obstruents in German*. Amsterdam: John Benjamins.
- Job, M. (1989) 'Sound change typology and the "Ejective Model"' In: Vennemann (1989a).
- Joos, M. (1952) 'The Medieval Sibilants' *Language* 28, 222-231. Reprinted in: Joos, M. (1958) *Readings in Linguistics*. New York: American Council of Learned Societies.
- Kager, R. (1999) *Optimality Theory*. Cambridge: CUP.
- Kahn, D. (1976) *Syllable-Based Generalizations in English Phonology*. PhD, MIT. Distributed by the IULC.
- Kaye, J. (1995) 'Derivations and interfaces' In: Durand & Katamba (1995)

- Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1985) 'The internal structure of phonological representations: a theory of charm and government' *Phonology Yearbook* 2, 305-328.
- Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1989) 'Rektion und Konstituentenrektion in der Phonologie' *Linguistische Berichte* 114, 31-75.
- Kaye, J., Lowenstamm, J. & Vergnaud, J.-R. (1990). 'Constituent structure and government in phonology' *Phonology* 7, 193-231.
- Keating, P. (1984) 'Phonetic and phonological representation of stop consonant voicing' *Language* 60, 286-319.
- Keller, R. (1978). *The German Language*. London: Faber.
- Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Oxford: Blackwell.
- Keyser, S & Stevens, K (1994) 'Feature geometry and the vocal tract' *Phonology* 11, 207-236.
- Kidd, D. (1957) *Latin Dictionary*. London: Collins.
- Kienle, R. (1960) *Historische Laut- und Formenlehre des Deutschen*. Tübingen: Niemeyer.
- King, R. (1969) *Historical Linguistics and Generative Grammar*. Englewood Cliffs: Prentice-Hall.
- Kiparsky, P. (1968) 'How abstract is phonology?'
- Kiparsky, P. (1979) 'Metrical structure assignment is cyclic' *Linguistic Inquiry* 10, 421-441.
- Kiparsky, P. (1982a) *Explanation in Phonology*. Dordrecht: Foris.
- Kiparsky, P. (1982b) 'Lexical morphology and phonology' In: Yang, I.-S. (ed) *Linguistics in the Morning Calm*. Hanshin: Seoul.
- Kiparsky, P. (1988) 'Phonological change' In: Newmeyer, F. (ed) *The Cambridge Linguistics Survey*. Volume 1. Cambridge: CUP.
- Kiparsky, P. (1995) 'The phonological basis of sound change' In: Goldsmith (1995).
- Kirchner, R. (1997) 'Contrastiveness and faithfulness' *Phonology* 14, 83-112.
- Kirchner, R. (1998) *An Effort-Based Approach to Consonant Lenition*. Unpublished PhD, University of California Los Angeles.
- Kirchner, R. (2000) 'Geminate inalterability and lenition' *Language* 76, 509-545.

- Knowles, G. (1974) *Scouse, the Urban Dialect of Liverpool*. Unpublished PhD dissertation, University of Leeds.
- Knowles, G. (1978). 'The nature of phonological variables in Scouse' In: Trudgill, P. (ed) *Sociolinguistic patterns in British English*. London: Edward Arnold.
- Koch, J. (1990) '**Cothairche*, Esposito's theory, and Neo-Celtic lenition' In: Bammesberger & Wollmann (1990).
- Koerner, K. (1989) 'Comments on reconstruction in historical linguistics.' In: Vennemann (1989).
- Kohler, K. (1977) *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt.
- Kohler, K. (1981) 'Three trends in phonetics: the development of the discipline in Germany since the nineteenth century' In: Asher, R. & Henderson, J, (eds) *Towards a History of Phonetics*. Edinburgh: EUP.
- Kohler, K. (1984) 'Phonetic explanation in phonology: the feature fortis/lenis' *Phonetica* 41, 150-174.
- Kohler, K. (1990) 'Comment on German' *Journal of The IPA* 20, 44-46.
- Kohler, K. (1995) *Phonetik des Deutschen*. Second Edition. Berlin: Erich Schmidt
- König, W. (1978) *Dtv-Atlas zur deutschen Sprache*. Munich: Deutscher Taschenbuch Verlag.
- Koopman, W., van der Leek, F., Fischer, O. & Eaton, R. (eds) (1987) *Explanation and Linguistic Change*. Amsterdam: Benjamins
- Kraemer, W. (1978) 'Umsetzungsschwierigkeiten von Dialektsprechern bei dem Gebrauch der Schriftsprache.' *Zeitschrift für Dialektologie und Linguistik Beiheft 27 Laut und Schrift in Dialekt und Standardsprache*, 93-219.
- Krahe, H. (1969) *Germanische Sprachwissenschaft I: Einleitung und Lautlehre*. Berlin: de Gruyter
- Kranzmayer, E. (1956) *Historische Lautgeographie des gesamtbairischen Dialektraumes*. Vienna: Böhlau.
- Kuhn, T. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Kuryłowicz, J. (1927) 'ə indo-européen et h en hittite' In: *Symbolae Grammaticae in Honorem Ioannis Rozwadowski*. Cracow: Gebethner & Wolff.
- Kuryłowicz, J. (1948) 'Le sens des mutations consonantiques' *Lingua* 1, 77-85.

- Labov, W. (1972) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1981) 'Resolving the neogrammarian controversy' *Language* 57, 267-308.
- Labov, W. (1994) *Principles of Linguistic Change*. Volume one: *Internal Factors* Oxford: Blackwell.
- Labov, W. (2001) *Principles of Linguistic Change*. Volume two: *Social Factors* Oxford: Blackwell.
- Ladefoged, P. (1971) *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Ladefoged, P. (1973) 'The features of the larynx' *Journal of Phonetics* 1, 73-83.
- Ladefoged, P. & Maddieson, I. (1996) *The Sounds of the World's Languages*. Oxford: Blackwell.
- Lass, R. (1974) 'Linguistic orthogenesis? Scots vowel length and the English length conspiracy' In: Anderson, J. & Jones, C. (eds.) *Historical Linguistics*. Volume 2: *Theory and Description in Phonology*. Amsterdam: North Holland.
- Lass, R. (1976) *English Phonology and Phonological Theory*. Cambridge: CUP.
- Lass, R. (1980) *On Explaining Language Change*. Cambridge: CUP.
- Lass, R. (1984) *Phonology*. Cambridge: CUP.
- Lass, R. (1984b) 'Reflections on rule loss' In: Balke, N. & Jones, C. (eds) *English Historical Linguistics: Studies in Development*. Sheffield: Centre for English Cultural Tradition and Language.
- Lass, R. (1991-1993) 'Old English fricative voicing revisited' *Studia Anglica Posnaniensia* 35-37, 3-46.
- Lass, R. (1992) 'Phonology and morphology' In: Blake, N. (ed) *The Cambridge History of the English Language*. Volume 2: 1066-1476. Cambridge: CUP.
- Lass, R. (1993) 'How real(ist) are reconstructions?' In: Jones, C. (ed) *Historical Linguistics: Problems and Perspectives*. London: Longman.
- Lass, R. (1994) *Old English*. Cambridge: CUP.
- Lass, R. (1997) *Language Change and Historical Linguistics*. Cambridge: CUP.
- Lass, R. & Anderson, J. (1975). *Old English Phonology*. Cambridge: CUP.

- Lavoie, L. (1996) 'Consonant strength: results of a data base development project' *Working Papers of the Cornell Phonetics Laboratory* 11, 269-316.
- Lavoie, L. (2000) *Phonological Patterns and Phonetic Manifestations of Consonantal Weakening*. PhD, Cornell University.
- Lavoie, L. (2001) *Consonantal Strength: Phonological Patterns and Phonetic Manifestations*. New York: Garland. A revised version of Lavoie (2000).
- Lehmann, W. (1973) *Historical Linguistics: an Introduction*. 2nd edition. New York: Holt, Rineheart & Winston.
- Lehmann, W. (1993) *Theoretical Bases of Indo-European Linguistics*. London: Routledge.
- Lessiak, P. (1933) *Beiträge zur Geschichte des deutschen Konsonantismus*. Brünn: Rohrer.
- Lightfoot, D. (1991) *How to set Parameters: Arguments from Language Change*. Cambridge, Mass: MIT Press
- Lightfoot, D. (1999) *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.
- Lisker, L & Abranson, A (1964) 'A cross-linguistic study of voicing in initial stops: acoustical measurements' *Word* 20, 384-422.
- Lockwood, W. (1969) *Indo-European Philology*. London: Hutchinson.
- Lombardi, L (1995) 'Laryngeal feature and privativity' *The Linguistic Review* 12, 35-60.
- Lombardi, L. (1990) 'The nonlinear organization of the affricate' *Natural Language and Linguistic Theory* 8, 375-426.
- Lombardi, L. (1991) *Laryngeal Features and Laryngeal Neutralisation*. PhD. University of Massachusetts at Amherst. Published (1994), New York: Garland.
- Luick, K. (1914-1940) *Historische Grammatik der englischen Sprache*. Reprinted (1964) Oxford: Blackwell.
- Macpherson, I (1975) *Spanish Phonology; Descriptive and Historical* Manchester: MUP.
- Maddieson, I. (1984) *Patterns of Sounds*. Cambridge: CUP.
- Maiden, M. (1995) *A Linguistic History of Italian*. London: Longman.
- Mangold, M. (1990) *Duden Ausspracheworterbuch: Worterbuch der deutschen Standardaussprache*. Mannheim: Bibliographisches Institut.

- Marchand, J. (1973) *The Sounds and Phonemes of Wulfila's Gothic*. The Hague: Mouton.
- Martinet, A. (1952) 'Celtic lenition and Western Romance consonants' *Language* 28, 192-217.
- Martinet, A. (1955) *Economie des changements phonétiques*. Berne: A Francke.
- Maza, M. (2000a) 'Phonetic data and functional explanation in phonology: the case of laxness vowel harmony in Granada Spanish' Newcastle: University of Newcastle upon Tyne, Manuscript.
- Maza, M. (2000b) 'Ambiguity and Laxness Vowel Harmony in Granada Spanish' Durham: University of Durham, Manuscript.
- McCarthy, J. (1979) *Formal Problems in Semitic Morphology and Phonology*. PhD, MIT. Published 1985. New York.: Garland.
- McCarthy, J. (1988) 'Feature geometry and dependency: a review' *Phonetica* 43, 84-108.
- McCarthy, J. (1989) 'Guttural Phonology' 2nd draft. Amherst, Mass.: University of Massachusetts, Manuscript.
- McCarthy, J. (2000) 'Sympathy and phonological opacity' *Phonology* 16, 331-400.
- McCarthy, J. & Prince, A. (1995) *Faithfulness and Reduplicative Identity*. Manuscript: University of Massachusetts, Amherst and Rutgers University.
- McMahon, A. (1984) *Understanding Language Change*. Cambridge: CUP.
- McMahon, A. (2000a) *Lexical Phonology and the History of English*. Cambridge: CUP
- McMahon, A. (2000b) *Change, Chance and Optimality*. Oxford: OUP.
- McMahon, A. (2001) 'Phonology and the Holy Grail' North West Centre for Linguistics Annual Lecture, University of Manchester.
- McMahon, A., Foulkes, P. & Tollfree, L. (1994) 'Gestural representation and Lexical Phonology' *Phonology* 11, 277-316.
- Meeussen, A. (1967) *Bantu Lexical Reconstructions*. Reprinted (1980) Tervuren: Koninklijk Museum voor Middenafrika.
- Meillet, A. (1937) *Introduction a l'étude comparative des langues indo-européennes*. Paris : Hachette.
- Meinhof, C. (1899) *Grundriß einer Lautlehre der Bantusprachen*. Berlin: Dietrich Reimer.

- Milroy, J. (1992) *Linguistic Variation and Change*. Oxford: Blackwell.
- Milroy, L. (1987) *Observing and Analysing Natural Language*. Oxford: Blackwell.
- Mohanan, K. (1986) *The Theory of Lexical Phonology*. Reidel: Dordrecht.
- Morris-Jones, J. (1913) *A Welsh Grammar*. Oxford: OUP.
- Mossé, F. (1952) *A Handbook of Middle English*. Baltimore: Johns Hopkins Press
- Moulton, W. (1954) 'The stops and spirants of early Germanic'. *Language* 30, 1-42.
- Nielsen, H. (1981) *Old English and the Continental Languages; a survey of Morphological and Phonological Interrelations*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Nurse, D. (1999) 'Towards a historical classification of East African Bantu languages' In: Hombert, J.-M., & Hyman, L. (eds) *Bantu Historical Linguistics*. Stanford: CSLI.
- Ó Lúing, S. (1991) *Kuno Meyer, 1858-1919: a Biography*. Dublin: Geography Publications.
- Ohala, J. (1974) 'Experimental historical phonology' In: Anderson, J. & Jones, C. (eds) *Historical Linguistics. Volume 2: Theory and Description in Phonology*. Amsterdam: North Holland.
- Ohala, J. (1987) 'Explanation, evidence, and experiment in phonology' In: Dressler, W, Luschützky, H., Pfeiffer, O & Rennison, J. (eds) *Phonologica 1984*. Cambridge: CUP.
- Ohala, J. (1992a) 'What's cognitive, what's not in, sound change.' In: Kellerman & Morrissey, (eds) *Diachrony within Synchrony*. Frankfurt: Peter Lang.
- Ohala, J. (1992b) 'The segment: primitive or derived?' In Docherty & Ladd (1992).
- Orton, H. (1962) *Survey of English Dialects: Introduction*. Leeds: E. J. Arnold.
- Palmer, L. (1980) *The Greek Language*. London: Faber & Faber.
- Pandeli, H., Eska, J., Ball, M. & Rahilly, J. (1997). 'Problems of phonetic transcription: The case of the Hiberno-English slit-t'. *Journal of the IPA* 27: 65-75.
- Passy P. (1890) *Étude sur les changements phonétiques*. Paris: Firmin-Didot.
- Paul, H. (1880) *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.
- Paul, H. (1916) *Deutsche Grammatik*. Volume 1. Halle: Niemeyer.

- Paul, H. (1944) *Mittelhochdeutschse Grammatik*. 14th edition, with E. Gierach. Halle: Niemeyer.
- Pedersen, H. (1897) *Aspirationen i Irsk; en sproghistorisk Undersøgelse*. Leipzig: M. Spirgatis.
- Pedersen, H. (1909) *Vergleichende Grammatik der keltischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht.
- Pedersen, H. (1951) *Die gemeinindoeuropäischen und die vorindoeuropäischen Verschußlaute*. Copenhagen: Munckgaard.
- Penny, R. (1991) *A History of the Spanish Language*. Cambridge: CUP.
- Penzl, H. (1970) 'Zu den Methoden der historischen Lautbestimmung: die ahd. Sibilanten'. *Folia Linguistica* 4, 104-109.
- Penzl, H. (1971) *Lautsystem und Lautwandel in den althochdeutschen Dialekten*. München: Max Hueber.
- Penzl, H. (1975) *Vom Urgermanischen zum Neuhochdeutschen*. Berlin: Erisch Schmidt.
- Peters, R. (1973) 'Mittelniederdeutscher Sprache.' In: Goossens, J. (ed) (1973) *Niederdeutsch: Sprache und Literatur: eine Einführung*. Volume 1: Sprache. Neumünster, 66-115.
- Picard, M. (1999) 'On spirantization and the minimality of phonological change' *Folia Linguistica Historica* 20, 63-77.
- Piggott, G. (1999) 'At the right edge of words' *The Linguistic Review* 16, 143-185.
- Polomé, E. (1965) 'The Laryngeal Theory so far'. In: Winter, W. (ed) *Evidence for Laryngeals*, The Hague: Mouton.
- Popper, K. (1959) *The Logic of Scientific Discovery*. London: Hutchinson.
- Popper, K. (1963) *Conjectures and Refutations*. London: Routledge and Kegan Paul.
- Posner, R. (1996) *The Romance Languages*. Cambridge: CUP.
- Postal, P. (1968) *Aspects of Phonological Theory*. New York: Harper & Row.
- Prince, A. (1984) 'Phonology with tiers'. In: Aronoff, M & Oehrle, R. (eds) *Language Sound Structure*. Cambridge, MA: MIT Press.
- Prince, A. & Smolensky, P. (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University & University of Colorado, Boulder, Manuscript.

- Prins, A. (1972) *A History of English Phonemes: from Indo-European to present-day English*. Leiden: Leiden University Press.
- Prokosch, E. (1933) *An Outline of German Historical Grammar*. New York: OUP.
- Prokosch, E. (1939) *A Comparative Germanic Grammar*. Philadelphia: LSA.
- Pullum, G. (1976) 'The Duke of York Gambit' *Journal of Linguistics* 12, 83-102.
- Ramat, A. (ed) (1998a) *The Indo-European languages*. London: Routledge
- Ramat, P. (1981) *Einführung in das Germanische*. Tübingen Niemeyer.
- Ramat, P (1998b) 'The Germanic Languages' In: Ramat (1998a).
- Rask, R. (1818) *Undersøgelse om det gamle Nordiske eller Islandske Sprogs Oprindelse*. Copenhagen: Gyldendalske.
- Raumer, R. von (1837) *Die Aspiration und die Lautverschiebung*. Reprinted (1972) Hildesheim: Gerstenberg.
- Rennison, J. (1986) 'On tridirectional feature systems for vowels' In: Durand (1986).
- Rennison, J. (1990) 'On the elements of phonological representations: the evidence from vowel systems and vowel processes' *Folia Linguistica* 24, 175-244.
- Robins, R. (1990) *A Short History of Linguistics*. London: Longman.
- Ruhlen, M. (1987) *A Guide to the World's Languages*. Stanford: Stanford University Press.
- Russ, C. (1982) *Studies in Historical German Phonology: a Phonological Comparison of MHG and NHG with Reference to Modern Dialects*. Berne: Peter Lang.
- Russ, C. (ed) (1990) *The Dialects of Modern German*. London: Routledge.
- Russell, P. (1995) *An Introduction to the Celtic Languages*. London: Longman.
- Sagey, E. (1986) *The Representation of Features and Relations in Nonlinear Phonology*. PhD, MIT. Published (1990) New York: Garland.
- Salmons, J. (1992) A look at the data for a global etymology: *tik 'finger'. In: Davis, G. & Iverson, G. (eds) *Explanation in Historical Linguistics*. Amsterdam: John Benjamins.
- Salmons, J. (1993) *The Glottalic Theory; Survey and Synthesis*. Journal of Indo-European Studies Monograph 10. Institute for the Study of Man .

- Sangster, C. (1999). *A study of lenition of /t/ and /d/ in Liverpool English*. Unpublished MPhil thesis, University of Oxford.
- Sangster, S. (to appear) 'Lenition of alveolar stops in Liverpool English' *Journal of Sociolinguistics*.
- Sapir, E. (1921) *Language*. New York: Harcourt Brace.
- Saussure, F. de (1879) *Mémoire sur le Système Primitif des Voyelles dans les Langues Indo-Européennes*. Leipzig: Vieweg.
- Saussure, F. de (1916) *Cours de Linguistique Générale*. Translated by Baskin, W. (1966) as *Course in General Linguistics*. New York: McGraw-Hill.
- Schane, S. (1984) 'The fundamentals of Particle Phonology' *Phonology Yearbook* 1, 129-155.
- Scheer, T. (1998) 'A theory of consonantal interaction' *Folia Linguistica* 32, 201-237.
- Schein, B. & Steriade, D. (1986) 'On Gemimates' *Linguistic Inquiry* 17, 691-744.
- Scherer, W. (1868) *Zur Geschichte der deutschen Sprache*. Berlin: Duncker.
- Schieb, G. (1970) 'Mittelhochdeutsch' In: Schmitt, L. (ed) *Kurzer Grundriß der germanischen Philologie bis 1500*. Berlin: de Gruyter, 347-385.
- Schirmunski, V. (1962) *Deutsche Mundartkunde*. Berlin: Akademie-Verlag.
- Schleicher, A (1861) *Compendium der vergleichende Grammatik der indogermanischen Sprachen*. Weimar: Böhlau.
- Schleicher, A. (1868) 'Eine fabel in indogermanischer ursprache'. *Beiträge zur vergleichenden Sprachforschung* 5, 205-208.
- Schmidt, O. & Vennemann, T. (1985a) 'Die Niederdeutschen Grundlagen des Standarddeutschen Lautsystems' Part 1. *Beiträge zur Geschichte der deutschen Sprache und Literatur*. 107, 1-20.
- Schmidt, O. & Vennemann, T. (1985b) 'Die Niederdeutschen Grundlagen des Standarddeutschen Lautsystems' Part 2. *Beiträge zur Geschichte der deutschen Sprache und Literatur*. 107, 157-173.
- Schmidt, W. (ed.) (1984) *Geschichte der deutschen Sprache*. Berlin: Volk und Wissen Volkseigener Verlag.
- Schrodt, R. (1976) *Die germanische Lautverschiebung und ihre Stellung im Kreise der indogermanischen Sprachen*. 2nd edition. Wien: Halosar.

- Schübel, G. (1955) *Die Ostfränkisch-Bambergische Mundart von Stadtsteinach im ehemaligen Fürstbistum Bamberg*. Giessen: Wilhelm Schmitz.
- Schwink, F. (1994) *Linguistic Typology, Universality and the Realism of Reconstruction*. Journal of Indo-European Studies Monograph 12. Institute for the Study of Man
- Selkirk, L. (1980) 'The role of prosodic categories in phonology' *Linguistic Inquiry* 11, 563-605.
- Sievers, E. (1876) *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf und Härtel.
- Sievers, E. (1901) *Grundzüge der Phonetik; zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Fifth edition. Leipzig: von Breitkopf & Härtel.
- Sims-Williams, N. (1990) 'Dating the transition to Neo-Brittonic: phonology and history 400-600' In: Bammesberger & Wollman (1990).
- Sims-Williams, N. (1998a) 'The Iranian languages' In: Ramat (1998).
- Sims-Williams, N. (1998b) 'The Celtic languages' In: Ramat (1998).
- Smith, N. (1981) 'Foley's scales of relative phonological strength' In: Goyvaerts (1981).
- Sommerstein, A. (1977) *Modern Phonology*. London: Edward Arnold.
- Sonderegger, S. (1979) *Grundzüge deutscher Sprachgeschichte*. Berlin: de Gruyter.
- Spencer, A. (1996) *Phonology*. Oxford: Blackwell.
- Stahlke, F. (1976) 'Segment sequences and segmental fusion' *Studies in African Linguistics* 7, 41-63.
- Stampe, D. (1973) 'On Chapter Nine' In: Kenstowicz, M. & Kisseberth, C. (eds) *Issues in Phonological Theory*. The Hague: Mouton.
- Stechow, A. von (1986) 'Notizen zu Vennemann's Anti-Grimm' *Beiträge zur Geschichte der deutschen Sprache und Literatur*. 108, 159-171.
- Steriade, D. (1995) 'Underspecification and markedness' In: Goldsmith (1995).
- Strang, B. (1970) *A History of the English Language*. London: Longman.
- Stroop, J. (1987) 'Explanation' by Linguistic Maps. In: Koopman, W., van der Leek, F., Fischer, O. & Eaton, R. *Explanation and Linguistic Change*. Amsterdam: Benjamins

- Stuart-Smith, J. (1999) 'Glasgow: accent and voice quality'. In: Foulkes & Docherty (1999).
- Sturtevant, E. (1917) *Linguistic change: an introduction to the historical study of language*. Chicago: University of Chicago Press.
- Sturtevant, E. (1940) *The Pronunciation of Greek and Latin*. 2nd edition. Philadelphia: LSA.
- Sweet, H. (1888) *A History of English Sounds from the Earliest Period with Full Word-Lists*. Oxford: Clarendon.
- Sweet, H. (1908) *The Sounds of English; an Introduction to Phonetics*. Oxford: Clarendon.
- Szemerényi, O. (1996) *Introduction to Indo-European linguistics*. Oxford: OUP.
- Thomas, P. (1990) 'The Brythonic consonant shift and the development of consonant mutation' *Bulletin of the Board of Celtic Studies* 37, 1-42.
- Thurneysen, R. (1898) Review of Pedersen (1897). *Anzeiger für indogermanische Sprach- und Altertumskunde. Beiblatt zu den Indogermanischen Forschungen* 9, 42-48.
- Thurneysen, R. (1909) *Handbuch des Alt-Irischen*. Part 1: Grammar.. Hedelberg: Carl Winter.
- Toorn, M. van den (1973) *Nederlandse Taalkunde*. Utrecht: Spectrum.
- Trask, R. L. (1996a) *Historical Linguistics*. London: Arnold.
- Trask, R. L. (1996b) *A Dictionary of Phonetics and Phonology*. London: Routledge.
- Trask, R. L. (1997) *The History of Basque*. London: Routledge.
- Trask, R. L. (2000) *The Dictionary of Historical and Comparative Linguistics*. Edinburgh: EUP.
- Trubetzkoy, N. (1939) *Grundzüge der Phonologie*. Translated by Baltaxe, C. (1969) as *Principles of Phonology*. Berkley: University of California Press.
- Tucker, A., & Bryan, M. (1957) *Linguistic Survey of the Northern Bantu Borderland*. London: OUP.
- Tuttle, E. (1916) *Dravidian Developments*. Language Monographs Series. Pennsylvania: Linguistic Society of America.
- Vaux, B. (1998) 'The laryngeal specifications of fricatives' *Linguistic Inquiry* 29, 497-511.
- Vekeman, H. & Ecke, A. (1992) *Geschichte der niederländischen Sprache*. Bern: Peter Lang.

- Vennemann, T. (1983) 'Causality in language change' *Folia Linguistica Historica* 6, 5-26.
- Vennemann, T. (1984) 'Hochgermanisch und Niedergermanisch; die Verzweigungstheorie der germanisch-deutschen Lautverschiebungen. *Beiträge zur Geschichte der deutschen Sprache und Literatur*. 106, 1-45.
- Vennemann, T. (1988) *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.
- Vennemann, T. (ed) (1989a) *The New Sound of Indo-European*. Berlin: Mouton de Gruyter.
- Vennemann, T. (1989b) 'Preface'. In: Vennemann (1989a).
- Vennemann, T. (1994) 'Dating the division between High and Low Germanic. A summary of the arguments' In: Swan, T. (ed.) *Older Germanic Languages in a comparative perspective*. Berlin: Mouton de Gruyter.
- Verner, K. (1875) 'Eine Ausnahme der ersten Lautverschiebung' *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 23, 97-130.
- Viëtor, W. (1893) 'Wie ist die Aussprache des Deutschen zu lehren? Ein Vortrag' Marburg/Lahn.
- Vijayakrishnan, K. (1999) 'Weakening Processes in the Optimality Framework' Paper presented at HILP 4 and stored at the Rutgers Optimality Archive.
- Vincent, N. (1988) 'Italian' In: Harris, M. & Vincent, N. (eds) *The Romance Languages*. London: Routledge.
- Wagner, M. (2000) 'Who's your next of kin?' Paper presented at the eighth Manchester Phonology Meeting, University of Manchester.
- Wang, W. (1969) 'Competing Changes as a Cause of Residue' *Language* 45, 9-25.
- Wang, W. & Cheng, C. (1977) 'Implementation of phonological change: the Shaungfeng Chinese case' In: Wang, W. (ed) *The Lexicon in Phonological Change*. The Hague: Mouton.
- Weinreich, U., Labov, W. & Herzog, M. (1968) 'Empirical foundations for a theory of language change' In: Lehmann, W. & Malkiel, Y. (eds) *Directions for Historical Linguistics*. Austin: University of Texas Press.
- Weldner, W. (1991) *Die Mundart von Barchfeld an der Werra. Zeitschrift für Dialektologie und Linguistik Beiheft 68*. Stuttgart: Steiner.
- Wells, J. (1982) *The Accents of English*. Three volumes. Cambridge: Cambridge University Press.

- Wetzels, L. & Sezer, E. (eds) (1986) *Studies in Compensatory Lengthening*. Dordrecht: Foris.
- Wiese, R. (1996) *The Phonology of German*. Oxford: OUP.
- Wireback, K. (1993) *The role of Phonological Structure in Sound Change from Latin to Spanish and Portuguese*. Unpublished PhD, the Pennsylvania State University.
- Wolf, N. (1981) *Geschichte der deutschen Sprache; Volume 1: Althochdeutsch, Mittelhochdeutsch*. Heidelberg: Quelle & Meyer.
- Wright, J. (1910) *Grammar of the Gothic Language*. Oxford: Clarendon.
- Wright, J. (1928) *An Elementary Middle English Grammar*. Oxford: OUP.
- Yoshida, S. (1996) *Phonological Government in Japanese*. Faculty of Asian Studies Monographs, new series 20. Canberra: Australian National University.
- Zabrocki (1965) 'Die dritte Lautverschiebung in Deutschen' *Symbolae Linguisticae in Honoram Georgii Kuryłowicz*. Cracow: Polska Akademia Nauk-Oddział w. Krokowie.
- Zehetner, L. (1977) *Bairisch. Dialekt/Hochsprache kontrastiv*. Volume 2. Düsseldorf: Padagogischer Verlag Schwann.
- Zvelebil, K. (1970) *Comparative Dravidian Phonology*. The Hague: Mouton.